Mutual Correlated Agreement and List-Decoding: Improvements and Separations

Ignacio Manzur
ignacio.manzur@epfl.ch
EPFL

November 3, 2026

Abstract

At a high level, our contribution is threefold: we improve the quantitative MCA error bound for polynomial generators, we prove a refinement of the implication from list-decodability to MCA, and we clarify that in general, these two notions are distinct.

We show that, for every finite field $\mathbb{F}$, every nonempty seed set $S \subseteq \mathbb{F}$, every linear code $\mathcal{C} \subseteq \mathbb{F}^n$ of relative minimum distance $\delta_{\mathcal{C}}$ and every degree $d \geq 1$, the univariate powers generator $G_d(x) := (1, x, \ldots, x^d)$ has mutual correlated agreement for $\mathcal{C}$ at relative proximity parameter γ with error $O(d \max\{1, n\gamma\}/|S|)$ below the smaller of the unique decoding radius and the generalized Johnson radius $J_{d+2}(\delta_{\mathcal{C}}) := 1 - (1 - \delta_{\mathcal{C}})^{1/(d+2)}$, with a one-coordinate margin below the latter when it is smaller.

Then, we show that list-decodability at radius p with list size L implies MCA slightly above $1 - \sqrt{1-p}$ (the improvement on the radius is dependent on the list size). The previous state-of-the-art result of Gao, Cai, Xu and Kan showed that list-decodability at radius $p < \delta_{\mathcal{C}}$ with list size L implies mutual correlated agreement at every proximity parameter strictly below $1 - \sqrt{1-p}$.

Finally, we construct two families of codes that show obstructions to any improvement of the $1 - \sqrt{1-p}$ bound, or to strong converse statements. Informally, a bound of the form $\epsilon_{\mathsf{MCA}}(1-(1-p)^{\alpha}) \leq Q(n, L)/|\mathbb{F}|$ for some polynomial Q and $\alpha > 1/2$ cannot hold for all linear codes $\mathcal{C}$ under the sole hypothesis that $\mathcal{C}$ is (p, L)-list-decodable. Similarly, for any distance $0 < \delta < 1$ and any $1 - \sqrt{1-\delta} < p < \gamma < 1$, there exists a family of linear codes with fixed positive rate and relative distance tending to δ such that $|\mathbb{F}|\, \epsilon_{\mathsf{MCA}}(\gamma) \leq \mathsf{poly}(n)$ but the list size at radius p is $n^{\omega(1)}$. Both counterexample families exist over fields of any characteristic. Both MCA errors are for the line generator.

Contents

1 Introduction **4**
- 1.1 Tight quantitative bounds for polynomial generators . . . 6
- 1.2 Mutual correlated agreement and list-decoding . . . 7
- 1.3 Related work . . . 9
- 1.4 Open questions . . . 10
- 1.5 AI assistance and author responsibility . . . 10

2 Technical overview **11**
- 2.1 Counting bad seeds on polynomial curves . . . 11
- 2.2 A list-size-dependent improvement to the MCA radius . . . 14
- 2.3 Bounded lists with many bad seeds . . . 17
- 2.4 Large lists with small MCA error . . . 19

3 Preliminaries **24**
- 3.1 Linear codes and list-decoding . . . 24
- 3.2 Generators and mutual correlated agreement . . . 25
- 3.3 Elementary bounds . . . 26

4 Bad seeds, curves, and root counting **27**

5 Agreement moments **32**
- 5.1 Moments of a family of sets . . . 32
- 5.2 Agreement sets of bad seeds . . . 35
- 5.3 List-decodability and the completion score . . . 35

6 List-decoding families of linear maps over prime fields **37**
- 6.1 A sunflower bound with repetitions . . . 37
- 6.2 Bounding the close rank-three maps . . . 37
- 6.3 The list bound . . . 39

7 A unified MCA bound for polynomial generators **40**

8 A list-size-dependent bound **44**
- 8.1 List-decodability bounds the completion score . . . 45
- 8.2 Distinct-slope sampling and the upper moment bound . . . 46
- 8.3 Proof of Theorem 8.1 . . . 47

9 An obstruction to uniform exponents above one half **48**
- 9.1 A trace subcode . . . 49
- 9.2 Puncturing while preserving list-decodability . . . 50
- 9.3 Proof of Corollary 9.2 . . . 52

10 Codes with large list size and small MCA error **53**
10.1 Generalized weights and mutual correlated agreement 54
10.2 Symmetric forms with large lists and large supports 58
10.3 Proof of Theorem 10.1 66
10.4 Proof of Corollary 10.2 72

A Numerical estimate for Lemma 4.11 **74**

B Numerical estimate for Theorem 7.1 **75**

C A simpler strong converse counterexample for binary fields **77**
C.1 Generalized Hamming weight amplification 77
C.2 Binary codes with large lists and large generalized weights 78
C.3 Distance and rate amplification 80
C.4 Proof of Theorem C.1 84

Acknowledgements **85**

References **85**

1 Introduction

We study the relationship between list-decoding and mutual correlated agreement (MCA) for linear codes. List-decoding is a classical and extensively studied notion in coding theory. MCA arose much more recently from a line of work on distance preservation and proximity gaps, motivated by the need for sound and efficient interactive oracle proofs (IOPs). Developments in the study of Reed–Solomon codes, which obtain correlated agreement and MCA at the same radius as list-decoding, and an implication from list-decodability to MCA with quadratic loss in the radius for arbitrary linear codes have brought these two notions increasingly close together, to the point where they are sometimes used interchangeably. Our main contribution is threefold: we improve the quantitative MCA error bound for polynomial generators, we prove a refinement of the implication from list-decodability to MCA, and we clarify that in general, these two notions are distinct. We recall their definitions and origins, and then introduce the questions that we study.

List-decoding. List-decoding [Eli57; Woz58] is a natural relaxation of unique decoding: instead of requiring a single codeword to be close to any word in Hamming distance, we allow a small list of close codewords. Let $\mathbb{F}$ be a finite field, let $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear code, and let Δ denote Hamming distance.

Definition 1. *For a relative radius $p \in [0,1]$ and an integer $L \geq 1$, the code $\mathcal{C}$ is* (p, L)-list-decodable *if every word $\mathbf{u} \in \mathbb{F}^n$ has at most L codewords within distance pn:*

$$|\{\mathbf{c} \in \mathcal{C} : \Delta(\mathbf{u}, \mathbf{c}) \leq pn\}| \leq L \ .$$

We write $L_{\mathcal{C}}(p)$ for the largest size of such a list over all words.

Unique decoding is the case $L = 1$. If $\mathcal{C}$ has relative minimum distance $\delta_{\mathcal{C}}$, then strictly below $\delta_{\mathcal{C}}/2$ every word has a unique close codeword. List-decoding allows us to consider larger radii, at the cost of considering a list of close codewords to a word.

A central question is the tradeoff between rate, decoding radius, and list size. For linear codes, a long line of work establishes capacity-approaching list-decodability, meaning the radius approaches $\delta_{\mathcal{C}}$, and improves the required list size [ZP81; GHSZ00; GHK10]. A major milestone is the construction of folded Reed–Solomon codes achieving list-decoding capacity [GR06].

Another fundamental question, more algorithmic in nature, asks for explicit codes whose decoding lists can also be found efficiently. Reed–Solomon codes have been a particularly important focus of this study. [Sud96] gave an efficient list-decoding algorithm beyond unique decoding for low-rate Reed–Solomon codes; [GS98] reached the Johnson radius. Subsequent work studies list-decodability beyond this radius for suitable evaluation sets [RW14; ST20; FKS22; GLSTW22]. [BGM23] prove that Reed–Solomon codes with randomly chosen evaluation points over exponentially large fields achieve list-decoding capacity with high probability. [GZ23] and [AGL24] obtain capacity results for randomly punctured Reed–Solomon codes over polynomial-size and linear-size fields, respectively. More recently, [BCPZZ26] give a polynomial-time list-decoding algorithm approaching capacity over prime fields on arbitrary evaluation sets.

Mutual correlated agreement. MCA comes from a different question: how can a verifier reduce testing the proximity of several words to a code to testing a single random combination? Such reductions play a central role in the soundness analysis of probabilistic proofs [RVW13; AHIV17]. The overarching notion is that of distance preservation: if one word is far away from the code, the combination of the words remains far except with small probability. The study of these reductions led to the definition of stronger forms of distance preservation, including correlated agreement [BCIKS20] and, recently, MCA [ACFY25], both of which additionally retain information about the coordinates on which the words agree with the code.

A *generator* $G\colon S \to \mathbb{F}^\ell$ maps a seed from an arbitrary domain to a vector of coefficients. Given words $\mathbf{u}_0, \ldots, \mathbf{u}_{\ell-1} \in \mathbb{F}^n$, it produces the combined word

$$\mathbf{u}_x := \sum_{j=0}^{\ell-1} G(x)_j \mathbf{u}_j \ .$$

MCA, introduced in [ACFY25], asks that large agreement sets of the combined word with the code should be explained by agreements of the original words *on the same coordinates*.

Definition 2 (Informal)**.** *For a relative radius $\gamma \in [0,1]$, the generator G has* mutual correlated agreement *for $\mathcal{C}$ with error $\epsilon_{\mathsf{MCA}}(\gamma)$ if, for every choice of the input words, all but an $\epsilon_{\mathsf{MCA}}(\gamma)$ fraction of seeds satisfy the following implication for every $T \subseteq [n]$:*

$$\begin{aligned} &|T| \geq (1-\gamma)n \text{ and } \mathbf{u}_x|_T \in \mathcal{C}|_T \\ &\implies \mathbf{u}_j|_T \in \mathcal{C}|_T \text{ for every } j \ . \end{aligned}$$

The exceptional seeds are called *MCA-bad seeds*. We use γ for the relative MCA radius and p for the relative list-decoding radius throughout the paper.

The parameters tracked in the MCA definition have direct consequences for the efficiency of proof systems based on proximity testing. A verifier can test a random combination of many committed words instead of testing each word separately. The MCA error ϵ_{MCA} contributes to the soundness error of this reduction, while the relative proximity parameter γ determines the distance available to eventual proximity tests. A larger usable proximity parameter can reduce the number of queries, and a smaller MCA error can reduce the field size or repetition needed for a target soundness error. In hash-based succinct arguments, fewer queries also reduce the cost of opening the committed words. These considerations motivate both the study of larger proximity parameters and sharper bounds on the error [ACFY24; ACFY25].

Connecting the two notions. Although list-decoding and MCA originated in different problems, their study has become closely connected. This is partly because of the situation for Reed–Solomon codes, where list-decoding techniques have also established correlated agreement and MCA at the same radii [BCIKS20; BCGM26; Jer26]. For general linear codes, [GCXK26] derive MCA from list-decodability, albeit with a quadratic loss in the radius. Together, these developments raise a natural question:

To what extent do list-decodability and small MCA error characterize the same codes?

Here, we should clarify that small MCA error is for the line generator $x \mapsto (1, x)$, whose simplicity often allows for a cleaner analysis, and whose MCA error is closely related to the error of the widely used random linear combination generator in any number of variables [BCGM26]. Maybe surprisingly, we show that these properties are substantially different for general linear codes. Our contribution is threefold:

1. We improve the quantitative MCA bounds for polynomial generators, matching the known lower-bound scaling over a larger range of proximity parameters within the unique decoding regime,

2. We sharpen the implication from list-decoding to MCA, obtaining an additive correction to the radius depending on the list size,

3. We construct linear codes over fields of any characteristic with bounded lists at radius p and large MCA error at any proximity parameter $1-(1-p)^\alpha$ for any $\alpha > 1/2$ (this shows that the exponent in the radius loss in the implication from list-decoding to MCA cannot be improved in general), and

we construct linear codes with small MCA error at proximity parameter γ but large lists at radius p for any $1 - \sqrt{1-\delta} < p < \gamma < 1$ and any relative distance $\delta \in (0,1)$ (this shows that a strong converse implication from MCA to list-decoding does not hold).

Together, these results roughly say that the MCA error for polynomial generators is optimal up to the unique decoding regime, and that our state of knowledge regarding the relationship between list-decodability and MCA is complete for general linear codes. We discuss these results in order.

1.1 Tight quantitative bounds for polynomial generators

A polynomial generator $G : S_1 \times \cdots \times S_s \subseteq \mathbb{F}^s \to \mathbb{F}^\ell$ is such that every one of its output coordinates is an s-variate polynomial in the input variables. A line of work on distance preservation and correlated agreement [RVW13; AHIV17; BCIKS20; BGKS20; GKL24; Hab25; Zei26] culminated in [BCGM26], which proves that all polynomial generators have MCA for every linear code, with explicit bounds on the radius and error. This essentially settled the question of which polynomial generators admit general MCA guarantees. It left a quantitative question about how closely the error can match the known lower bound.

Consider the univariate powers generator

$$\begin{aligned} G_d &\colon S \subseteq \mathbb{F} \to \mathbb{F}^{d+1} \ , \\ G_d(x) &= (1, x, \ldots, x^d) \ . \end{aligned}$$

For this generator, [BCGM26] gives an error bound of

$$\max\{1, n\gamma\} \cdot \frac{d}{|S|}$$

up to radius $\delta_{\mathcal{C}}/(d+2)$. It also gives a lower bound of order $n\gamma d/\left|\mathbb{F}\right|$ throughout the unique decoding regime when $S = \mathbb{F}$, when d divides $|\mathbb{F}| - 1$ and $|\mathbb{F}| \geq n\gamma d + 1$. Beyond $\delta_{\mathcal{C}}/(d+2)$, its upper bound reaches the generalized Johnson radius

$$J_{d+2}(\delta_{\mathcal{C}}) := 1 - (1 - \delta_{\mathcal{C}})^{1/(d+2)} \ ,$$

but incurs a multiplicative loss depending on the slack from that radius. The remaining quantitative question is how far the lower-bound scaling $n\gamma d/|S|$ can be matched within the unique decoding regime.

Our first result matches this scaling over a larger range of proximity parameters, with constants independent of the degree. We distinguish the two possible orderings of the unique decoding and generalized Johnson radius.

Theorem 1 (Informal)**.** *Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear code of relative distance $\delta_{\mathcal{C}}$, let $d \geq 1$, and let $S \subseteq \mathbb{F}$ be a nonempty seed set. For the univariate powers generator $G_d(x) = (1, x, \ldots, x^d)$, the following bounds hold:*

1. *If $\delta_{\mathcal{C}}/2 \leq J_{d+2}(\delta_{\mathcal{C}})$, then for every $0 \leq \gamma < \delta_{\mathcal{C}}/2$,*

$$\epsilon_{\mathsf{MCA}}(\gamma) \leq \frac{24d}{|S|} \max\{1, n\gamma\} \ . \tag{1}$$

2. *If $J_{d+2}(\delta_{\mathcal{C}}) < \delta_{\mathcal{C}}/2$, then for every $0 \leq \gamma \leq J_{d+2}(\delta_{\mathcal{C}}) - 1/n$,*

$$\epsilon_{\mathsf{MCA}}(\gamma) \leq \frac{15d}{|S|} \max\{1, n\gamma\} \ .$$

See Theorem 7.1 and Corollary 7.2 for the proofs and the rounding convention for arbitrary real proximity parameters.

1.2 Mutual correlated agreement and list-decoding

For general linear codes, the only known implication from list-decodability to MCA is proved in [GCXK26]. If a code is (p, L)-list-decodable with $p < \delta_{\mathcal{C}}$, this result gives MCA for affine combinations at proximity parameters strictly below

$$1 - \sqrt{1-p} \ ,$$

with an error polynomial in the blocklength, list size, and inverse slack, divided by the field size. The loss in the radius is substantial: list-decodability at p does not yield MCA at p by this argument.

The situation for Reed–Solomon codes has been markedly different. At the Johnson radius, the list-decoding and correlated-agreement guarantees reach the same threshold [BCIKS20], and subsequent work establishes MCA there as well [BCHKS26; BCGM26]. [GG26] and [GGSW26] obtain near-capacity proximity gaps and MCA for subspace-design codes and suitable random ensembles, including Reed–Solomon codes with random evaluation points. Most recently, [BCPZZ26] prove list-decoding up to capacity over prime fields on arbitrary evaluation sets. Building on the techniques of [BCPZZ26], [Jer26] proves MCA at the same fixed-slack radii, with polynomially many exceptional parameters and no proximity loss. [HKS26] give a unified exposition of these developments.

These results suggest a close connection between list-decoding and MCA: in the Reed–Solomon settings where the strongest bounds are known, the two guarantees reach the same radius. But this connection uses additional structure of the codes. We show that *list-decodability and small MCA error are not equivalent properties of general linear codes*.

An improved implication from list-decoding. We first sharpen the implication of [GCXK26] for the line generator, with an additive correction to the proximity parameter depending on the list size.

Theorem 2 (Informal)**.** *Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear, (p, L)-list-decodable code with $|\mathbb{F}| \geq 2L$. For $0 \leq \gamma \leq p < 1$, define*

$$F_{L+1}(\gamma) := (1-\gamma)^2 + \frac{\gamma + (1-\gamma)\gamma^{L+1}}{L+1} \ .$$

If $G := F_{L+1}(\gamma) - (1-p) > 0$, then the line generator $x \mapsto (1, x)$ has MCA for $\mathcal{C}$ with error

$$\epsilon_{\mathsf{MCA}}(\gamma) = O\left(\frac{L(n\gamma+1)}{G\,|\mathbb{F}|}\right) \ .$$

For a fixed positive margin G, our bound is linear in L, improving the quadratic dependence on L in [GCXK26]. The positive correction to $(1-\gamma)^2$ gives radii strictly above $1 - \sqrt{1-p}$. It vanishes as L grows, recovering the square-root threshold. See Theorem 8.1 for the explicit bound, which also applies whenever $|\mathbb{F}| > L$, and for the integer-radius conventions.

An obstruction to improving the exponent. Despite this finite-list improvement, the exponent $1/2$ cannot be increased uniformly under a list-decodability hypothesis alone, even in a prescribed characteristic.

Theorem 3 (Informal)**.** *Fix a prime ℓ and an exponent $\alpha > 1/2$. There is no polynomial $Q \in \mathbb{R}[X, Y, Z]$ for which the bound*

$$\epsilon_{\mathsf{MCA}}(\gamma) \leq \frac{Q(n, L, s^{-1})}{|\mathbb{F}|} \ ,$$
$$s := 1 - (1-p)^{\alpha} - \gamma > 0 \ ,$$

holds for all linear, (p, L)-list-decodable codes over fields $\mathbb{F}$ of characteristic ℓ. The MCA error is for the line generator. This remains true if the coefficients of Q depend on ℓ, α, p and γ. Writing $\theta := \min\{\alpha - 1/2, 1/2\}$, the witnessing codes can be constructed over $\mathbb{F}_{\ell^m}$ for every $m \geq m_0$, where

$$\begin{aligned} m_0 &= O_\ell\left(\log \frac{2}{\theta}\right), \\ L &= O_\ell\left(\theta^{-(2\ell^6+2)}\right), \\ n &= O_\ell\left(\theta^{-(2\ell^6+4)} m^2\right) \quad . \end{aligned}$$

For every fixed prime ℓ and sufficiently small fixed $\eta > 0$, our codes have bounded list size $L_{\ell,\eta}$ at radius $p = 1 - \ell^{-2} - \eta$ and MCA error at least $(1 - 1/|\mathbb{F}|)/(\ell - 1)$ at radius $\gamma = 1 - \ell^{-1} + \eta$; see Theorem 9.1 and Corollary 9.2. The list size is independent of $|\mathbb{F}|$, while the MCA error stays bounded away from zero as $|\mathbb{F}| = \ell^m$ grows. In particular, even bounded list size does not generally imply small MCA error at the same radius.

This is consistent with the finite-list improvement in Theorem 2. The list bound $L_{\ell,\eta}$ depends on η, and its correction to the square-root threshold is too small to reach the radius $1 - \ell^{-1} + \eta$. At these parameters, $F_{L_{\ell,\eta}+1}(\gamma) < 1 - p$, so the positive-gap hypothesis fails; see (35). Thus a list-dependent improvement in the radius can coexist with this construction.

Small MCA error does not imply small lists. We also rule out a converse implication: a code with distance $\delta_{\mathcal{C}}$ can have small MCA error at γ and still have superpolynomial lists at a radius p, for $1 - \sqrt{1 - \delta_{\mathcal{C}}} < p < \gamma < 1$. We clarify again that the MCA error is for the line generator.

This separation is surprising because a large decoding list provides many codewords close to the same word, each a potential codeword witness for MCA-badness. Since $p < \gamma$, these codewords are close even at a radius strictly smaller than the MCA proximity parameter. One might therefore expect a large list to supply witnesses for many bad seeds. This intuition may also help explain the expectation that list-decodability and small MCA error should be closely related.

For these counterexamples with superpolynomial lists, we focus on radii beyond the Johnson radius. Below it, the Johnson bound already gives polynomial list size whenever the slack is at least inverse-polynomial in the blocklength, independently of any MCA assumption.

Theorem 4 (Informal)**.** *Fix a prime ℓ, a relative distance $0 < \delta_{\mathcal{C}} < 1$, and*

$$1 - \sqrt{1 - \delta_{\mathcal{C}}} < p < \gamma < 1 \quad .$$

There is a family of linear codes $\mathcal{C} \subseteq \mathbb{F}^n$ over fields of characteristic ℓ, with rate tending to a fixed positive constant and relative distance tending to $\delta_{\mathcal{C}}$, such that

$$\begin{aligned} L_{\mathcal{C}}(p) &= n^{\omega(1)} \quad , \\ \epsilon^{\mathcal{C}}_{\mathsf{MCA}}(\gamma) &\leq \frac{C_0 n^k}{|\mathbb{F}|} \quad , \end{aligned}$$

where MCA is for the line generator and C_0, k depend only on $\ell, \delta_{\mathcal{C}}, p, \gamma$. To quantify the size requirements, put

$$\xi := \min\{\delta_{\mathcal{C}}, 1 - \delta_{\mathcal{C}}, p - J_2(\delta_{\mathcal{C}}), (1 - \gamma)^3\} > 0 \quad .$$

The codes are scalar extensions of codes over a fixed field of size $q = O_\ell(\xi^{-2})$. The bounds hold along the family for every $n \geq n_0$, with

$$k, n_0 \leq \exp\left(\exp\left(O_\ell\left(\xi^{-1}\log^2\frac{2}{\xi}\right)\right)\right) \quad .$$

A field size $|\mathbb{F}| = n^{O(k)}$ suffices to make the MCA error at most n^{-1}.

The polynomial field size specified above makes the MCA error tend to zero while the lists remain superpolynomial. Both properties hold for the same codes; see Theorems C.1 and 10.1.

1.3 Related work

Distance-based counterexamples. [KTZ26] construct linear codes at every relative distance $\delta_{\mathcal{C}}$ for which the line generator has a constant fraction of MCA-bad seeds beyond the cubic radius $J_3(\delta_{\mathcal{C}}) = 1-(1-\delta_{\mathcal{C}})^{1/3}$. This limits the MCA radius guaranteed by minimum distance alone. This result is complementary to our quantitative bound, which improves the error below the generalized Johnson radius, within the unique decoding regime.

Reed–Solomon counterexamples over binary fields. [DKTZ26] construct Reed–Solomon codes over binary fields with superpolynomial decoding lists and MCA error exceeding any uniform $\mathsf{poly}(n)/|\mathbb{F}|$ bound beyond the Johnson radius. Our obstruction has bounded list size at a larger radius and blocklength polylogarithmic in the field size, ruling out a stronger general implication from list-decodability to MCA.

From large lists to related codes with large MCA error. [GYXK26] transform a code with a large decoding list into a related code with large MCA error by changing at most one generator-matrix column, preserving length and dimension and decreasing minimum distance by at most one. Our converse separation shows that the original code can have small MCA error even at a radius larger than the list-decoding radius.

Ingredients of our constructions. The trace-polynomial mechanism behind our first obstruction already appears in DEEP-FRI [BGKS20]. We combine it with a bounded-list result for families of linear maps over any fixed prime field at rank distance at least three, and puncture to polylogarithmic blocklength while preserving the required distances and list bound. The reduction to low-rank differences is related to [GGR09]; the puncturing argument uses the ordinary-to-average-radius conversion and plurality methods of [RW14; RW15].

For the converse separation, we use generalized Hamming weights [Wei91] to bound MCA independently of list size. Generalized weights already appear in list-decoding bounds in [GGR09], and shortening dimensions are used to obtain MCA error bounds in [Cho26b]. Our contribution is a criterion using a single higher generalized weight that applies uniformly to scalar extensions: counting short equations over the fixed base field keeps the exceptional-seed bound independent of the extension-field size. We give the precise comparisons in Section 10.1.

Codes defined by symmetric multilinear forms over any fixed prime field have large decoding lists while their high-dimensional subspaces have large support, which gives small MCA error after scalar extension. Related precedents include an improved weight bound for symmetric matrices with zero diagonal in a tensor-product code [AK11] and codes obtained by partially evaluating symmetric multilinear forms at rows of an MDS generator matrix [AK24]. A simpler construction from binary BCH codes appears in Appendix C.

1.4 Open questions

Our two counterexample families exist in every characteristic, but both use nontrivial extensions of a fixed prime field. What happens over prime fields themselves, with their order growing with the blocklength? Can list-decodability imply MCA beyond the square-root threshold in this setting, or can the obstruction in Theorem 3 be realized over prime fields? In the reverse direction, can prime-field codes have fixed positive rate, superpolynomial decoding lists, and small MCA error at a larger radius, as in Theorem 4, or does restricting to prime fields permit a positive converse?

The situation for Reed–Solomon codes is nothing short of mysterious. The known results suggest that list-decodability and MCA are equivalent in this case, but the techniques used to prove this equivalence are highly specialized. A very interesting question would be to uncover a fundamental property of Reed–Solomon codes that explains this equivalence, and to determine whether it extends to other families of linear codes.

1.5 AI assistance and author responsibility

The author relied on AI interaction, specifically GPT-6.1 Sol, to refine and sometimes discover the results in this paper. The author's original proof of the uniform MCA error scaling was refined by GPT-6.1 Sol until the agreement moment technique surfaced, which the author realized could also be applied to obtain the list-size improvement to the MCA radius. The trace code mechanism is an example from DEEP-FRI [BGKS20], and it followed that if the code was also list-decodable and could be punctured it would immediately give a counterexample to improvements to the $1/2$ exponent from [GCXK26], the proof of which GPT-6.1 Sol provided. To control MCA error independently of list-size, GPT-6.1 Sol provided a beautiful argument using a single generalized Hamming weight [Wei91] to count short equations over the fixed base field uniformly across scalar extensions (see Section 10.1 for related precedents), and constructed a family of multilinear form codes simultaneously having large lists and large generalized Hamming weights for any characteristic. The author verified the correctness of all statements and proofs, and takes full responsibility for any errors or omissions.

2 Technical overview

We first explain the quantitative bound for polynomial generators, then the improved transfer from list-decoding. After that, we describe the two constructions that separate list-decodability from MCA.

2.1 Counting bad seeds on polynomial curves

We review the ideas leading to Theorem 1, our quantitative MCA bound for polynomial generators. Fix words $\mathbf{U} \coloneqq \mathsf{Mat}(\mathbf{u}_0, \ldots, \mathbf{u}_d)$, a proximity parameter $\gamma \in [0, 1]$, and the polynomial word

$$\mathbf{u}(\mathtt{X}) = \sum_{j=0}^{d} \mathbf{u}_j \mathtt{X}^j = G_d(\mathtt{X}) \cdot \mathbf{U} \ ,$$

where $G_d(\mathtt{X})$ is the degree d univariate powers generator $\mathtt{X} \mapsto (1, \mathtt{X}, \mathtt{X}^2, \ldots, \mathtt{X}^d)$. For a coordinate $i \in [n]$, we denote its scalar polynomial by

$$\mathbf{u}_i(\mathtt{X}) := \sum_{j=0}^{d} \mathbf{u}_j[i] \mathtt{X}^j \ .$$

Here, $\mathbf{u}_j[i]$ denotes the i-th coordinate of the word $\mathbf{u}_j$. More generally, $P_i(\mathtt{X})$ denotes the i-th coordinate polynomial of a vector-valued polynomial $P(\mathtt{X})$. Thus j indexes the input words, while i indexes their coordinates. We call a seed x *bad* if there exists a set $T_x \subseteq [n]$ with $|T_x| \geq (1 - \gamma)n$ and $\mathbf{c}_x \in \mathcal{C}$ such that $\mathbf{u}(x)|_{T_x} = \mathbf{c}_x|_{T_x}$, but there exists some j such that $\mathbf{u}_j|_{T_x} \notin \mathcal{C}|_{T_x}$. Within the unique decoding radius, the codeword is unique, and we may take T_x to be its full agreement set with $\mathbf{u}(x)$. Our goal is to bound the number of points $(x, \mathbf{c}_x)$. To do so, we split the bad seeds into two classes: those whose associated codewords lie on a single polynomial curve with coefficients in $\mathcal{C}$, and those that do not. This strategy [GKL24; Hab25; BCGM26; Zei26] is useful as the size of the former can already be controlled by a bound of the form $O(n\gamma d)$, where the hidden constant depends on a careful choice of a threshold for the size of the common agreement set of the associated codewords. We bound the size of the latter by introducing the notion of *agreement moments* (a version of which appeared in DEEP-FRI [BGKS20]), which allows us to obtain a bound of the form $O(n\gamma d)$ on the remaining seeds.

Interpolation and root counting. The associated codewords at any $d + 1$ distinct bad seeds determine a unique curve $P(\mathtt{X}) = \sum_{j=0}^{d} \mathbf{c}_j^* \mathtt{X}^j$ of degree at most d, by interpolation. Its coefficients $\mathbf{c}_j^*$ are linear combinations of the witness codewords, so they belong to $\mathcal{C}$. On the intersection of the agreement sets of these bad seeds, P and $\mathbf{u}$ agree identically: each coordinate difference has $d + 1$ distinct roots and degree at most d.

We call a tuple of $d + 2$ distinct seeds *coherent*[1] if all its associated codewords lie on one such curve, and *noncoherent* otherwise. For a noncoherent tuple, fix an ordering of the seeds and let x be the $(d + 2)$-th seed. Its associated codeword $\mathbf{c}_x$ does not lie on the curve P through the first $d + 1$ codewords. The nonzero codeword $\mathbf{c}_x - P(x)$ nevertheless vanishes on the common agreement set of all $d + 2$ bad seeds, since both codewords agree there with $\mathbf{u}(x)$. Minimum distance therefore bounds the common agreement fraction of a noncoherent tuple by $1 - \delta_{\mathcal{C}}$. This is the baseline for our upper bound on the expected common agreement of a random tuple: only coherent tuples can contribute more than $1 - \delta_{\mathcal{C}}$. Controlling their frequency will let us

[1] Coherence is our local terminology for the configurations sought in curve-decodability [GG26; GGSW26]. Our proof bounds their frequency; Remark 4.8 gives the precise connection with curve-decodability and curve-free subspaces.

compare this upper bound with the lower bound forced by the fact that the agreement sets T_x are large for a bad seed x.

Root counting also bounds how many bad seeds can have their associated codewords on a single curve P. Write T_P for the coordinates where P and $\mathbf{u}$ agree identically: equality of the polynomial coefficients gives $\mathbf{u}_j|_{T_P} = \mathbf{c}_j^*|_{T_P}$ for every j. For a seed with $\mathbf{c}_x = P(x)$, we have $T_P \subseteq T_x$. If no coordinate outside T_P contributed an additional agreement, then $T_x = T_P$, and each original word $\mathbf{u}_j$ would agree on all of T_x with the codeword $\mathbf{c}_j^*$. This would contradict the seed being bad. Thus every bad seed on P is a root of at least one nonzero coordinate polynomial $\mathbf{u}_i(\mathtt{X}) - P_i(\mathtt{X})$ with $i \notin T_P$. Moreover, the condition $\Delta(\mathbf{u}(x), \mathbf{c}_x) \leq \gamma n$ requires at least $n - |T_P| - \gamma n$ of these nonzero coordinate polynomials to vanish at x. Each has at most d roots. Counting these incidences, together with the requirement of at least one root per bad seed, gives at most $d(n\gamma + 1)$ bad seeds on any one curve. If $|T_P|/n$ is bounded away from $1 - \gamma$, the same count gives a sharper bound independent of n; see Lemma 4.5.

Removing large curves and comparing adjacent moments. Let B be the set of bad seeds, that is:

$$\mathsf{B} := \left\{ x \in S : \begin{array}{l} \exists\, T \subseteq [n] \text{ s.t.} \\ \quad |T| \geq n \cdot (1-\gamma) \\ \quad \wedge\, (G_d(x) \cdot \mathbf{U})|_T \in \mathcal{C}|_T \\ \quad \wedge\, \exists\, j \in \{0, \ldots, d\}, \mathbf{u}_j|_T \notin \mathcal{C}|_T \end{array} \right\} .$$

The case $|\mathsf{B}| \leq d(n\gamma + 1)$ gives the desired bound directly. Assuming otherwise, we set a threshold τ satisfying $1 - \delta_{\mathcal{C}} < \tau < 1 - \gamma$, and call a curve through $d + 1$ associated codewords *large* if $|T_P|/n > \tau$. Split B into the seeds whose associated codewords lie on at least one large curve and the remaining set B', of size M. We can show that for a certain choice of τ, there are at most 4 large curves (see Lemma 4.11). Root counting bounds the number of bad seeds on each by $d(n\gamma + 1)$, giving

$$|\mathsf{B}| \leq 4d(n\gamma + 1) + M \ .$$

It remains to bound M. If $M \leq d + 1$, its contribution is already $O(d)$. Otherwise, sample $d + 1$ distinct seeds $x_1, \ldots, x_{d+1}$ uniformly from B' and let P be the curve through their associated codewords. Their common agreement set is exactly T_P. Now sample one more seed y uniformly from the remaining $M - d - 1$ seeds. Since P is not large, $|T_P|/n \leq \tau < 1 - \gamma$, and therefore every bad seed x_j, for $j \in [d + 1]$, must be the root of some coordinate polynomial $\mathbf{u}_i(\mathtt{X}) - P_i(\mathtt{X})$, for some $i \notin T_P$. The sharper root-counting bound of Lemma 4.5 gives at most $d + d\gamma/(1 - \gamma - \tau)$ bad seeds on P in total. The $d + 1$ interpolation seeds are already among them, leaving at most $d\gamma/(1 - \gamma - \tau) - 1$ additional seeds. Dividing by the $M - d - 1$ possible choices of y gives

$$\Pr\left[\mathbf{c}_y = P(y) \mid x_1, \ldots, x_{d+1}\right] \leq \frac{d\gamma}{(M - d - 1)(1 - \gamma - \tau)} \ .$$

Let Coh denote the event that $\mathbf{c}_y = P(y)$, conditioned on the first $d + 1$ seeds being fixed, so that P and T_P are fixed. All expectations and probabilities below are over the remaining seed y. Splitting according to coherence gives

$$\begin{aligned} &\mathbb{E}_y\left[\frac{|T_P \cap T_y|}{n}\right] \\ &= \mathbb{E}_y\left[\frac{|T_P \cap T_y|}{n} \cdot \mathbb{1}_{\mathsf{Coh}}\right] + \mathbb{E}_y\left[\frac{|T_P \cap T_y|}{n} \cdot \mathbb{1}_{\neg\mathsf{Coh}}\right] \end{aligned}$$

$$
\begin{aligned}
&\leq \Pr_y[\mathsf{Coh}] \cdot \frac{|T_P|}{n} + \Pr_y[\neg\mathsf{Coh}] \cdot (1-\delta_{\mathcal{C}}) \\
&= (1-\delta_{\mathcal{C}}) + \Pr_y[\mathsf{Coh}] \cdot \left(\frac{|T_P|}{n} - (1-\delta_{\mathcal{C}}) \right) \\
&\leq (1-\delta_{\mathcal{C}}) + \frac{d\gamma}{(M-d-1)(1-\gamma-\tau)} \cdot \max\left\{0, \frac{|T_P|}{n} - (1-\delta_{\mathcal{C}})\right\} \ .
\end{aligned}
$$

The first contribution comes from noncoherent tuples, whose common agreement fraction is at most $1-\delta_{\mathcal{C}}$. The second comes from coherent tuples: the codeword associated to the bad seed y lies on P, so $T_P \subseteq T_y$ and the intersection remains T_P. The last inequality uses the root-counting probability bound and discards a negative excess if $|T_P|/n < 1-\delta_{\mathcal{C}}$.

Our threshold satisfies $1-\delta_{\mathcal{C}} < \tau$, and $|T_P|/n \leq \tau$. On this interval, the positive-part function lies below the line joining its values at 0 and τ, giving

$$
\begin{aligned}
&\max\left\{0, \frac{|T_P|}{n} - (1-\delta_{\mathcal{C}})\right\} \\
&\leq \frac{\tau - (1-\delta_{\mathcal{C}})}{\tau} \cdot \frac{|T_P|}{n} \ .
\end{aligned}
$$

Thus the conditional upper bound depends linearly on $\frac{|T_P|}{n}$, which is the quantity we will average next. To average this conditional bound over the first $d+1$ seeds, we need to keep track of the expected sizes of two intersections: the intersection before adding the last seed, and the intersection after adding it. This motivates the introduction of the *agreement moments*. For $1 \leq j \leq M$, let μ_j be the expected common agreement fraction of j distinct uniformly sampled seeds. If coordinate i occurs in $A_i \in \mathbb{N}$ agreement sets, then

$$
\begin{aligned}
\mu_j &:= \mathbb{E}\left[\frac{|\bigcap_{r=1}^{j} T_{x_r}|}{n}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} \frac{A_i(A_i-1)\cdots(A_i-j+1)}{M(M-1)\cdots(M-j+1)} \ .
\end{aligned}
$$

In particular, averaging $|T_P|/n$ gives μ_{d+1}, while averaging the conditional agreement after adding y gives μ_{d+2}. We therefore obtain:

$$
\mu_{d+2} \leq (1-\delta_{\mathcal{C}}) + \frac{d\gamma}{M-d-1} \cdot \frac{\tau-(1-\delta_{\mathcal{C}})}{\tau(1-\gamma-\tau)} \cdot \mu_{d+1} \ .
$$

We are consequently led to compare these two adjacent moments. They cannot vary independently since both are determined by the same coordinate frequencies A_i. Convexity gives the complementary inequality

$$
\mu_{d+2} \geq \frac{M\mu_{d+1}^{(d+2)/(d+1)} - (d+1)\mu_{d+1}}{M-d-1} \ .
$$

Moreover, every agreement set has size at least $(1-\gamma)n$, which forces $\mu_{d+1}^{1/(d+1)} \geq 1-\gamma-O(d\gamma/M)$. If M were arbitrarily large, the lower bound on μ_{d+2} would approach $(1-\gamma)^{d+2}$, while the upper bound would approach $1-\delta_{\mathcal{C}}$ because coherent tuples become rare. Below the generalized Johnson radius,

$(1-\gamma)^{d+2} > 1-\delta_{\mathcal{C}}$, so these bounds are incompatible for sufficiently large M. This contradiction bounds M. Comparing the precise inequalities exposes the gap

$$1-\gamma-(1-\delta_{\mathcal{C}})^{1/(d+2)} = J_{d+2}(\delta_{\mathcal{C}})-\gamma \ ,$$

and bounds M by a quantity of the form $O(d\max\{1,\gamma/(J_{d+2}(\delta_{\mathcal{C}})-\gamma)\})$. Adding the removed seeds gives (1). Keeping these two contributions separate avoids multiplying the entire blocklength-dependent bound by the inverse gap. The details appear in Sections 4, 5 and 7.

2.2 A list-size-dependent improvement to the MCA radius

We review the ideas leading to Theorem 2, which improves the MCA radius in the implication from list-decoding to MCA with an additive term depending on the list size. Suppose that $\mathcal{C} \subseteq \mathbb{F}^n$ is (p, L)-list-decodable with $|\mathbb{F}| > L$, and consider the line $\mathbf{u}(x) = \mathbf{v} + x\mathbf{w}$, where $\gamma \leq p$. For a bad seed x, we no longer have unique witness codewords $\mathbf{c}_x$, but we can fix one once and for all and trim its agreement set T_x to size $(1-\gamma)n$.

Removing large codeword lines. The argument starts similarly as in the previous section by removing large codeword lines. For a codeword line $Q(\mathtt{X}) = \mathbf{q}_0 + \mathtt{X}\mathbf{q}_1$ with $\mathbf{q}_0, \mathbf{q}_1 \in \mathcal{C}$, let

$$T_Q := \{i \in [n] : Q_i(\mathtt{X}) \equiv \mathbf{u}_i(\mathtt{X})\} \ .$$

We call Q *large* if $|T_Q| \geq (1-p)n$. The threshold is the agreement required for list-decoding at radius p: on T_Q, the codeword $Q(x)$ agrees with $\mathbf{u}(x)$ for every seed, so a large line satisfies

$$\Delta(Q(x), \mathbf{u}(x)) \leq n - |T_Q| \leq pn \text{ for every } x \in \mathbb{F} \ .$$

Consequently, at each seed the evaluations of all large lines take at most L distinct values. Distinct codeword lines can have the same evaluation at no more than one seed. Counting these collisions over $\mathbb{F}$ bounds the number N of large lines by

$$\begin{aligned} |\mathbb{F}|\left(\frac{N^2}{L} - N\right) &\leq N(N-1) \\ \implies N &\leq \frac{L(|\mathbb{F}|-1)}{|\mathbb{F}|-L} \ . \end{aligned}$$

See Lemma 4.13 for the counting argument. Again let B denote the set of bad seeds for $\mathbf{u}$, that is

$$\mathsf{B} := \left\{ x \in \mathbb{F} : \begin{array}{c} \exists T \subseteq [n] \text{ s.t.} \\ |T| \geq n \cdot (1-\gamma) \\ \wedge\, \mathbf{v} + x\mathbf{w}|_T \in \mathcal{C}|_T \\ \wedge\, (\mathbf{v}|_T \notin \mathcal{C}|_T \text{ or } \mathbf{w}|_T \notin \mathcal{C}|_T) \end{array} \right\} \ .$$

Remove every bad seed whose chosen witness point $(x, \mathbf{c}_x)$ lies on one of the large lines. Root counting bounds the number of such seeds on each line by $\gamma n + 1$, so the remaining set B' satisfies

$$|\mathsf{B} \setminus \mathsf{B}'| \leq \frac{L(|\mathbb{F}|-1)}{|\mathbb{F}|-L}(\gamma n + 1) \ .$$

For every $x \neq y$ in B', the line Q through $(x, \mathbf{c}_x)$ and $(y, \mathbf{c}_y)$ now has $|T_Q| < (1-p)n$.

A fixed seed and its slopes. Let $M := |\mathsf{B}'|$. Fix a seed $x \in \mathsf{B}'$. A seed y in $\mathsf{B}' \setminus \{x\}$ defines the codeword slope

$$d_{x,y} := \frac{\mathbf{c}_y - \mathbf{c}_x}{y - x} \in \mathcal{C} \ .$$

On a coordinate $i \in T_x \cap T_y$, we have $\mathbf{c}_y[i] - \mathbf{c}_x[i] = (y - x)\mathbf{w}[i]$, so $d_{x,y}[i] = \mathbf{w}[i]$. Thus each seed y in $\mathsf{B}' \setminus \{x\}$ supplies a codeword agreeing with $\mathbf{w}$ over $T_x \cap T_y$. If M is too large, our goal is to choose $L + 1$ distinct slopes and construct a single word within relative distance p of all of them, contradicting list-decodability. The precise condition for this construction to be impossible will provide both the desired bound on M and the improved MCA radius.

Completing the agreements. First fix seeds $y_1, \ldots, y_{L+1} \in \mathsf{B}' \setminus \{x\}$ with pairwise distinct slopes, and consider $\ell_j := |T_x \cap T_{y_j}|$. Since the corresponding codeword line was not removed as large, $\ell_j < (1 - p)n$. To reach the agreement threshold within the list-decodability radius, slope j therefore needs at least

$$e_j := (1 - p)n - \ell_j > 0$$

additional agreements. Recall that each slope d_{x,y_j} agrees with $\mathbf{w}$ on $T_x \cap T_{y_j}$. Every coordinate outside $\bigcup_j (T_x \cap T_{y_j})$ is available for what we call completion: we may set its value to that of one chosen slope without destroying any of the agreements. The available coordinates are exactly

$$\mathcal{D} := ([n] \setminus T_x) \cup \left(T_x \setminus \bigcup_{j=1}^{L+1} T_{y_j} \right) \ ,$$
$$|\mathcal{D}| = \gamma n + H \ ,$$

where $H := |T_x \setminus \bigcup_j T_{y_j}|$ counts coordinates in T_x that none of the neighbors uses. If $\sum_j e_j \le |\mathcal{D}|$, choose pairwise disjoint subsets $E_j \subseteq \mathcal{D}$ with $|E_j| = e_j$. Define a word $\mathbf{z}$ to equal $\mathbf{w}$ on $\bigcup_j (T_x \cap T_{y_j})$, to equal d_{x,y_j} on E_j, and to be arbitrary elsewhere. The sets E_j are disjoint, so these assignments are consistent. Each slope agrees with $\mathbf{z}$ on its original overlap $T_x \cap T_{y_j}$ and on its assigned completion set E_j, giving

$$\Delta(\mathbf{z}, d_{x,y_j}) \le n - \ell_j - e_j = pn \ .$$

The total number of additional agreements required is

$$\sum_{j=1}^{L+1} e_j = (L + 1)(1 - p)n - \sum_{j=1}^{L+1} \ell_j \ .$$

Thus the completion condition $\sum_j e_j \le |\mathcal{D}|$ can be rewritten as

$$\sum_{j=1}^{L+1} \ell_j + \gamma n + H \ge (L + 1)(1 - p)n \ .$$

This motivates tracking the left-hand side, normalized by $(L + 1)n$, as the completion score

$$Z(x; y_1, \ldots, y_{L+1}) := \frac{\sum_{j=1}^{L+1} \ell_j + \gamma n + H}{(L + 1)n} \ .$$

The normalized target is $1-p$: if $Z \geq 1-p$, the completion procedure above produces a word within radius p of all $L+1$ slopes. List-decodability rules this out, so every tuple with distinct slopes must satisfy $Z < 1-p$.

Averaging the completion budget. To find tuples with a large score, sample a seed $\mathcal{X}$ uniformly from B' and then sample $Y_1, \ldots, Y_{L+1}$ independently from $\mathsf{B}' \setminus \{\mathcal{X}\}$. For the moment, we allow the slopes to repeat; we will address repetitions below. The score asks us to average two quantities: the size of the overlaps $|T_{\mathcal{X}} \cap T_{Y_j}|$, and $|T_{\mathcal{X}} \setminus \bigcup_{j=1}^{L+1} T_{Y_j}|$. The expected overlap sizes are equal for all j. This naturally leads to the definition of an agreement-disagreement moment ν_{L+1}:

$$\begin{aligned}
\hat{\mu}_2 &:= \mathbb{E}\left[\frac{|T_{\mathcal{X}} \cap T_{Y_1}|}{n}\right] \ , \\
\nu_{L+1} &:= \mathbb{E}\left[\frac{|T_{\mathcal{X}} \setminus \bigcup_{j=1}^{L+1} T_{Y_j}|}{n}\right] \ , \\
\mathbb{E}[Z] &= \hat{\mu}_2 + \frac{\gamma}{L+1} + \frac{\nu_{L+1}}{L+1} \ .
\end{aligned}$$

The moment $\hat{\mu}_2$ is the same as the agreement moment defined in the previous section, except that every trimmed agreement set has size $(1-\gamma)n$. As before, convexity gives a lower bound on $\hat{\mu}_2$. For a coordinate belonging to $T_{\mathcal{X}}$, independence makes the probability that all neighbors miss it the $(L+1)$-st power of the probability that one neighbor misses it. Averaging these probabilities and applying convexity relates ν_{L+1} to $\hat{\mu}_2$:

$$\begin{aligned}
\hat{\mu}_2 &\geq (1-\gamma)^2 - \frac{(1-\gamma)\gamma}{M-1} \ , \\
\nu_{L+1} &\geq (1-\gamma)\left(1 - \frac{\hat{\mu}_2}{1-\gamma}\right)^{L+1} \ .
\end{aligned}$$

Together they yield

$$\begin{aligned}
\mathbb{E}[Z] &\geq (1-\gamma)^2 + \frac{\gamma + (1-\gamma)\gamma^{L+1}}{L+1} - \frac{(1-\gamma)\gamma}{M-1} \\
&= F_{L+1}(\gamma) - \frac{(1-\gamma)\gamma}{M-1} \ .
\end{aligned}$$

Here, we have used $F_{L+1}(\gamma) := (1-\gamma)^2 + \frac{\gamma+(1-\gamma)\gamma^{L+1}}{L+1}$. The pairwise term alone suggests the square-root threshold $(1-\gamma)^2 > 1-p$. Keeping the completion terms gives the positive finite-list correction in F_{L+1}.

The size of this correction matters for the obstruction below. It is at most $1/(L+1)$, since $\gamma + (1-\gamma)\gamma^{L+1} \leq 1$. A family with a list bound independent of the field size can therefore fall outside the range $F_{L+1}(\gamma) > 1-p$ if that list bound is large relative to the gap between $(1-\gamma)^2$ and $1-p$. This is exactly what happens for the family in Section 9; the comparison is made explicit in (35).

Ensuring distinct slopes. The lower bound uses independent neighbors, while the list-decoding contradiction requires distinct slopes. Fix the seed x. Seeds in $\mathsf{B}' \setminus \{x\}$ with the same slope d lie on the same codeword line $Q(\mathtt{X}) = \mathbf{c}_x + (\mathtt{X} - x)d$, together with x itself. Since this line is not large, its deficit $D := n - |T_Q|$ satisfies $D \geq pn + 1$. The sharper inequality in Lemma 4.12 bounds the number of bad seeds on Q by $D/(D-\gamma n)$. Subtracting the fixed seed, the number of neighbors sharing slope d is therefore at most

$$\frac{D}{D-\gamma n} - 1 = \frac{\gamma n}{D-\gamma n} \leq R := \frac{\gamma n}{n(p-\gamma)+1} \ .$$

If $M-1 \le LR$, this already bounds the remaining set. Otherwise, any L slope classes contain at most $LR < M-1$ neighbors, so there must be at least $L+1$ classes. We can then repair the independent sample one at a time. At step $j \in [L+1]$, keep the sampled seed unless its slope repeats one of the previously retained slopes; in that case, replace it by a seed from an unused slope class. The $j-1$ used classes contain at most $(j-1)R$ of the $M-1$ seeds in $\mathsf{B}' \setminus \{x\}$, so the probability of replacement is at most $(j-1)R/(M-1)$.

In the numerator of Z, coordinates outside $T_{\mathcal{X}}$ contribute only through the fixed term γn, so replacing a neighbor does not change their contribution. For $i \in T_{\mathcal{X}}$, let k_i be the number of neighbors whose agreement sets contain i. This coordinate contributes $k_i + \mathbb{1}_{\{k_i=0\}}$ to $\sum_j \ell_j + H$. Replacing one neighbor changes k_i by at most one, and hence changes this contribution by at most one. Since $|T_{\mathcal{X}}| = (1-\gamma)n$, the normalized score changes by at most $(1-\gamma)/(L+1)$. Let $\widetilde{Z}$ be the score after these replacements. The repaired slopes are distinct, so $\widetilde{Z} < 1-p$, while the expected replacement cost gives

$$\mathbb{E}[|Z-\widetilde{Z}|] \le \frac{1-\gamma}{L+1}\sum_{j=1}^{L+1}\frac{(j-1)R}{M-1} = \frac{(1-\gamma)LR}{2(M-1)} \ ,$$
$$\mathbb{E}[Z] < 1-p+\frac{(1-\gamma)LR}{2(M-1)} \ .$$

Comparing this upper bound with the lower bound on the score, and writing $G := F_{L+1}(\gamma) - (1-p) > 0$, bounds M in both cases:

$$M \le 1 + \max\left\{LR, \frac{(1-\gamma)\gamma + (1-\gamma)LR/2}{G}\right\} \ .$$

Finally, add the seeds removed on large codeword lines and divide by $|\mathbb{F}|$ to obtain Theorem 8.1. See Sections 5.3, 8.1 and 8.2 for the detailed moment and coupling arguments.

2.3 Bounded lists with many bad seeds

We review the ideas leading to Theorem 3, our obstruction to increasing the square-root exponent in the list-decoding to MCA implication. We extend the trace-polynomial example from DEEP-FRI [BGKS20] to every characteristic. Fix a prime ℓ, let $|\mathbb{F}| = \ell^m$ with $m \ge 3$, and consider

$$\mathcal{C}_0 := \left\{ z \longmapsto \sum_{j=2}^{m-1} \alpha_j z^{\ell^j} : \alpha_j \in \mathbb{F} \right\} \ ,$$

evaluated on all of $\mathbb{F}$, together with $\mathbf{u}(z) := z$ and $\mathbf{v}(z) := z^\ell$. The code omits the two lowest Frobenius powers, z and z^ℓ. This makes $\mathbf{u}$ at relative distance at least $1-\ell^{-2}$ from the code, while many combinations $\mathbf{u} + x\mathbf{v}$ are within relative distance $1-\ell^{-1}$.

The distance of the input and its combinations. The Frobenius map $\zeta \mapsto \zeta^{|\mathbb{F}|/\ell^2}$ permutes $\mathbb{F}$. Reindexing the coordinates by $z = \zeta^{|\mathbb{F}|/\ell^2}$ turns a codeword into

$$\mathbf{c}(\zeta^{|\mathbb{F}|/\ell^2}) = \sum_{j=2}^{m-1} \alpha_j \zeta^{\ell^{j-2}} \ ,$$

a polynomial of degree at most $|\mathbb{F}|/\ell^3$. Here we use $\zeta^{\ell^m} = \zeta$. Under the same permutation, $\mathbf{u}$ becomes the monomial $\zeta^{|\mathbb{F}|/\ell^2}$. Its difference from any codeword has degree exactly $|\mathbb{F}|/\ell^2$, so it can vanish on at most that many coordinates. This gives the lower bound $1-\ell^{-2}$ on the relative distance of $\mathbf{u}$ from the code.

For the combinations, the useful seeds are the nonzero $(\ell-1)$-st powers

$$\mathcal{P} := \{a^{\ell-1} : a \in \mathbb{F}^\times\} \ , \qquad |\mathcal{P}| = \frac{|\mathbb{F}|-1}{\ell-1} \ .$$

For $x = a^{\ell-1}$, expansion of the trace gives a codeword witness

$$\mathbf{c}_x(z) := -\sum_{j=2}^{m-1} x^{(\ell^j-1)/(\ell-1)} z^{\ell^j} \in \mathcal{C}_0 \ ,$$
$$\mathbf{u}(z) + x\mathbf{v}(z) - \mathbf{c}_x(z) = \frac{\mathsf{Tr}_{\mathbb{F}/\mathbb{F}_\ell}(az)}{a} \ .$$

The trace functional is nonzero and has a kernel of size $|\mathbb{F}|/\ell$. Thus the combination agrees with its witness on an ℓ^{-1} fraction of the coordinates. On that same set, $\mathbf{u}$ cannot agree with any codeword, since this would put it within relative distance $1-\ell^{-1}$ of the code. Every seed in $\mathcal{P}$ is therefore MCA-bad at radius $1-\ell^{-1}$.

From rank distance to bounded lists. Each codeword is an $\mathbb{F}_\ell$-linear map. The difference of two distinct codewords has at most $|\mathbb{F}|/\ell^3 = \ell^{m-3}$ zeros by the degree bound above. Rank-nullity therefore gives

$$|\ker(\mathbf{c}-\mathbf{c}')| \le \ell^{m-3} \quad \Longrightarrow \quad \mathsf{rank}(\mathbf{c}-\mathbf{c}') \ge 3 \ .$$

To bound the decoding lists, we use the fact that rank-three maps at pairwise rank distance at least three have bounded lists at every agreement above ℓ^{-2}. More precisely, at most $D_{\ell,\eta}$ such maps can agree with any word on an $\ell^{-2}+\eta$ fraction of the coordinates, where the bound depends only on ℓ, η; see Lemma 6.2 for the proof.

We apply this estimate through a graph on the codewords in the decoding list of a word $f\colon \mathbb{F} \to \mathbb{F}$. Join two codewords when their difference has rank exactly three. Subtracting a fixed vertex from its neighbors gives a family to which the rank-three estimate applies, so every degree is at most $D_{\ell,\eta}$. Adjacent codewords have agreement ℓ^{-3}, and nonadjacent codewords have agreement at most ℓ^{-4}. Suppose the list contains N codewords, each agreeing with f on at least an $\ell^{-2}+\eta$ fraction. For a uniformly sampled coordinate $z \in \mathbb{F}$, let $A(z)$ count the codewords in the list agreeing with $f(z)$. Then $\mathbb{E}[A] \ge N(\ell^{-2}+\eta)$, so $\mathbb{E}[A^2] \ge N^2(\ell^{-2}+\eta)^2$. On the other hand, $A(z)(A(z)-1)$ counts ordered pairs of distinct codewords in the list both agreeing with $f(z)$, and hence with each other. At most $ND_{\ell,\eta}$ ordered pairs are adjacent in the graph; bounding their agreement by ℓ^{-3} and all other pairwise agreements by ℓ^{-4}, and using $\mathbb{E}[A] \le N$, gives

$$N^2(\ell^{-2}+\eta)^2 - N \le \frac{N(N-1) + (\ell-1)ND_{\ell,\eta}}{\ell^4} \ .$$

Since $(\ell^{-2}+\eta)^2 > \ell^{-4}$, this bounds N independently of m. Thus $\mathcal{C}_0$ has bounded lists at radius $1-\ell^{-2}-\eta$; see Proposition 6.3.

Keeping the blocklength small. The full evaluation code has length $|\mathbb{F}|$, which is too large to rule out a polynomial numerator in the MCA bound. We therefore sample $n = O_{\ell,\eta}((\log|\mathbb{F}|)^2)$ coordinates uniformly and independently, restricting the code and both input words to this sample. Concentration preserves the distances between $\mathbf{u}$ and every codeword, and between each combination indexed by $\mathcal{P}$ and its chosen witness. A union bound makes all these estimates hold simultaneously. The punctured $\mathbf{u}$ stays at relative distance at least $1-\ell^{-2}-\eta$ from the code, while each selected combination stays within relative distance $1-\ell^{-1}+\eta$ of its witness.

To preserve bounded lists, we must rule out a word agreeing with many codewords on the sampled coordinates, including words that depend on the sample. Fix r distinct original codewords $\mathbf{c}_1, \ldots, \mathbf{c}_r$. At coordinate z, the largest fraction of them that any one symbol can match is

$$Y(z) := \frac{1}{r} \max_{a \in \mathbb{F}} |\{j \in [r] : \mathbf{c}_j(z) = a\}| \ .$$

Let $g(z)$ be a symbol attaining this maximum. Then $\mathbb{E}_z[Y(z)]$ is exactly the average, over $j \in [r]$, of the agreement fraction between g and $\mathbf{c}_j$. The full code has list size at most some constant $B = B_{\ell,\eta}$ at agreement $\ell^{-2} + \eta/4$. Thus at most B of these codewords can agree with g on more than that fraction; each of the others contributes at most $\ell^{-2} + \eta/4$ to the average. Choosing a fixed $r \geq 4B/\eta$ gives

$$\mathbb{E}_z[Y(z)] \leq \ell^{-2} + \eta/4 + B/r \leq \ell^{-2} + \eta/2 \ .$$

For sampled coordinates $z_1, \ldots, z_n$, concentration and a union bound over all choices of r codewords ensure that

$$\frac{1}{n} \sum_{i=1}^{n} Y(z_i) < \ell^{-2} + \eta$$

for every such choice. If a word on the sampled coordinates agreed with each of r restricted codewords on at least an $\ell^{-2} + \eta$ fraction, their average agreement with that word would be at least this threshold. But at each sampled coordinate, that word can match at most a $Y(z_i)$ fraction of them, contradicting the displayed inequality. Hence every decoding list of the restricted code has size less than r, independently of m; see Section 9.2.

Choose $0 < \eta < (\ell^{-1} - \ell^{-2})/2$, so that

$$p := 1 - \ell^{-2} - \eta > \gamma := 1 - \ell^{-1} + \eta \ .$$

For every $x \in \mathcal{P}$, the punctured combination agrees with its witness on at least an $\ell^{-1} - \eta$ fraction of the coordinates. The distance lower bound on $\mathbf{u}$ prevents those agreements from extending it to a codeword. Hence every seed in $\mathcal{P}$ remains MCA-bad, and

$$\epsilon^{\mathcal{C}}_{\mathsf{MCA}}(\gamma) \geq \frac{1 - 1/|\mathbb{F}|}{\ell - 1} \ .$$

Finally, fix $\alpha > 1/2$. Since $\ell^{-2\alpha} < \ell^{-1}$, we can choose a sufficiently small fixed η so that

$$s := 1 - (1-p)^{\alpha} - \gamma = \ell^{-1} - \eta - (\ell^{-2} + \eta)^{\alpha} > 0 \ .$$

Along the resulting family, $\ell, p, \gamma, s, L_{\ell,\eta}$ are fixed and the blocklength is quadratic in $\log |\mathbb{F}|$. Every proposed bound $Q(n, L_{\ell,\eta}, s^{-1})/|\mathbb{F}|$ therefore tends to zero, while the MCA error stays bounded away from zero. This rules out the stronger exponent in every prescribed characteristic, even when the polynomial's coefficients depend on the characteristic, p and γ; see Theorem 9.1 and Corollary 9.2.

2.4 Large lists with small MCA error

We review the ideas leading to Theorem 4, our separation between small MCA error and small decoding lists. We work over a finite extension $\mathbb{K}/\mathbb{F}_q$, with $|\mathbb{K}| = q^e$ and $e = [\mathbb{K} : \mathbb{F}_q]$. The base field size $q = \ell^m$ remains fixed as $|\mathbb{K}|$ grows, and its characteristic may be any prescribed prime ℓ. A natural strong converse

would assert that a polynomial bound on $|\mathbb{K}|\epsilon_{\mathsf{MCA}}^{\mathcal{C}_{\mathbb{K}}}(\gamma)$ forces a polynomial bound on the decoding list size at a fixed smaller radius $p < \gamma$. We rule out such an implication even for codes of fixed positive rate and relative distance tending to any prescribed $0 < \delta_{\mathcal{C}} < 1$. More precisely, for every pair of fixed radii

$$1 - \sqrt{1 - \delta_{\mathcal{C}}} < p < \gamma < 1 \ ,$$

we construct codes $\mathcal{C}_n \subseteq \mathbb{F}_q^n$ at an unbounded sequence of lengths n, such that, for every finite extension $\mathbb{K}/\mathbb{F}_q$, their scalar extensions $\mathcal{C}_{n,\mathbb{K}} := \mathcal{C}_n \otimes_{\mathbb{F}_q} \mathbb{K}$ satisfy

$$\begin{aligned} L_{\mathcal{C}_{n,\mathbb{K}}}(p) &= n^{\omega(1)} \ , \\ |\mathbb{K}|\epsilon_{\mathsf{MCA}}^{\mathcal{C}_{n,\mathbb{K}}}(\gamma) &\leq C_0 n^k \ , \end{aligned}$$

as $n \to \infty$ along this sequence, with C_0, k independent of the blocklength and extension field size. Thus even allowing a fixed loss in radius does not give a converse from small MCA error to small lists in this regime; see Theorems C.1 and 10.1.

This is surprising because every codeword within relative distance p of a word is also within relative distance γ of it. A superpolynomial list therefore supplies many potential codeword witnesses already at a radius strictly smaller than the MCA radius. One might expect this abundance of proximate codewords to force the MCA error beyond $\mathsf{poly}(n)/|\mathbb{K}|$. Nevertheless, our codes have small MCA error despite these large lists, and choosing $|\mathbb{K}| = n^{O(k)}$ as in Theorem 4 makes the error tend to zero while their lists remain superpolynomial.

The separation uses a large *generalized Hamming weight* [Wei91]. For a code $\mathcal{C} \subseteq \mathbb{F}_q^n$, its k-th generalized weight $d_k(\mathcal{C})$ is the smallest size of the union of supports of a k-dimensional subcode. Thus every subcode supported on fewer than $d_k(\mathcal{C})$ coordinates has dimension below k. We prove that, for fixed q, k, the condition

$$\frac{d_k(\mathcal{C})}{n} > 1 - (1 - \gamma)^3$$

with a fixed positive margin gives MCA error $\mathsf{poly}(n)/|\mathbb{K}|$ for every scalar extension $\mathcal{C}_{\mathbb{K}} = \mathcal{C} \otimes_{\mathbb{F}_q} \mathbb{K}$. We construct codes satisfying this condition while having superpolynomial decoding lists at radius p. Generalized weights and subcode dimensions have also been used in list-decoding [GGR09] and MCA bounds [Cho26b]; precise comparisons appear in Section 10.1.

Localizing the MCA-bad property with two fixed seeds. Fix words $\mathbf{u}, \mathbf{v} \in \mathbb{K}^n$. For each MCA-bad seed x, choose a witness codeword $\mathbf{c}_x \in \mathcal{C}_{\mathbb{K}}$ and an agreement set T_x, and write $E_x := [n] \setminus T_x$. Thus $|E_x| \leq \gamma n$, and $\mathbf{u} + x\mathbf{v}$ agrees with $\mathbf{c}_x$ outside E_x, but at least one word from $\mathbf{u}, \mathbf{v}$ does not agree with the code over those same coordinates.

Fix two distinct bad seeds x_0, x_1. Their witness codewords determine a unique affine codeword line $Q_0(\mathtt{X})$ with $Q_0(x_0) = \mathbf{c}_{x_0}$ and $Q_0(x_1) = \mathbf{c}_{x_1}$. Subtract this line and write

$$\begin{aligned} f(\mathtt{X}) &:= \mathbf{u} + \mathtt{X}\mathbf{v} - Q_0(\mathtt{X}) = \mathbf{a} + \mathtt{X}\mathbf{b} \ , \\ S &:= E_{x_0} \cup E_{x_1} \ . \end{aligned}$$

At every coordinate outside S, the residual polynomial $f(\mathtt{X})$ vanishes at x_0 and x_1. Since it is affine, it vanishes identically there. Hence both residual words $\mathbf{a}, \mathbf{b}$ are supported on S. Write $Q_0(\mathtt{X}) = \mathbf{q}_0 + \mathtt{X}\mathbf{q}_1$, where $\mathbf{q}_0, \mathbf{q}_1 \in \mathcal{C}_{\mathbb{K}}$. On each witness set T_x, the residual combination $\mathbf{a} + x\mathbf{b}$ agrees with the codeword $\mathbf{c}_x - Q_0(x)$. Moreover, $\mathbf{a} = \mathbf{u} - \mathbf{q}_0$ agrees with a codeword on T_x if and only if $\mathbf{u}$ does, and the same holds for $\mathbf{b} = \mathbf{v} - \mathbf{q}_1$ and $\mathbf{v}$. Thus at least one of the residual words still fails to agree with any codeword on T_x.

Subtracting Q_0 therefore localizes the problem to the fixed coordinate set S. Now consider a third bad seed x_2, with error set E_{x_2}, and put

$$\begin{aligned} U &:= S \cup E_{x_2} \ , \\ J &:= S \cap T_{x_2} = S \setminus E_{x_2} \ . \end{aligned}$$

The residual witness $\mathbf{c}_{x_2} - Q_0(x_2)$ is supported on U: outside this union, both $f(x_2)$ and its error vanish. The coordinates on which it must match $f(x_2)$ inside S are precisely J. Let

$$W := \{\mathbf{c}|_J : \mathbf{c} \in \mathcal{C}_{\mathbb{K}},\ \mathsf{supp}(\mathbf{c}) \subseteq U\} \ .$$

Let w be a word supported on S. To find a codeword equal to w on T_{x_2}, we need to match w on $J = S \cap T_{x_2}$ and set the codeword to zero on $T_{x_2} \setminus S$, where w is already zero. Since $T_{x_2} \setminus S = [n] \setminus U$, this is exactly the requirement that the codeword be supported on U and restrict to w on J. By definition of W, such a codeword exists if and only if $w|_J \in W$.

The word $\mathbf{a} + x_2\mathbf{b}$ equals the codeword $\mathbf{c}_{x_2} - Q_0(x_2)$ on T_{x_2}, whereas at least one of $\mathbf{a}, \mathbf{b}$ does not equal any codeword on that set. Applying the equivalence above to these three words gives

$$\begin{aligned} &(\mathbf{a} + x_2\mathbf{b})|_J \in W \ , \\ &\mathbf{a}|_J \notin W \text{ or } \mathbf{b}|_J \notin W \ . \end{aligned}$$

If $|U| < d_k(\mathcal{C})$, the space of codewords supported on U has dimension at most $k-1$, and therefore so does W. This dimension bound gives short equations for membership in W. Choose an information set $I \subseteq J$ of size $\dim W \leq k-1$: the coordinates in I determine every word in W, and each other coordinate is a linear combination of them. Since W is obtained from the base code by shortening, restriction and scalar extension, these combinations have coefficients in $\mathbb{F}_q$. Thus membership is expressed by equations of the form

$$\begin{aligned} &w_j - \sum_{i \in I} \alpha_{j,i} w_i = 0 \text{ for } j \in J \setminus I \ , \\ &\alpha_{j,i} \in \mathbb{F}_q \ . \end{aligned}$$

Each equation involves at most $|I| + 1 \leq k$ coordinates. Extending the coefficient vector of such an equation by zero outside J gives $\lambda \in \mathbb{F}_q^n$ with

$$\begin{aligned} &|\mathsf{supp}(\lambda)| \leq k \ , \\ &\lambda \cdot \mathbf{a} + x_2 \lambda \cdot \mathbf{b} = 0 \ , \\ &(\lambda \cdot \mathbf{a}, \lambda \cdot \mathbf{b}) \neq (0,0) \ . \end{aligned}$$

For fixed x_0, x_1, the words $\mathbf{a}$ and $\mathbf{b}$ are fixed. The last condition makes the preceding affine equation in x_2 nonzero, so each coefficient vector λ can account for at most one third seed. Its entries lie in the fixed base field $\mathbb{F}_q$, even though the seed and residual word entries lie in $\mathbb{K}$.

Normalize the first nonzero entry of λ to one. The number of possible normalized equations is

$$\mathsf{T}_k(n,q) := \sum_{j=1}^{k} \binom{n}{j} (q-1)^{j-1} \ .$$

Although E_{x_2}, U, J and W depend on the third seed, every resulting equation belongs to this same collection. We can therefore count these equations instead of the possible agreement sets: for fixed x_0, x_1, at most $\mathsf{T}_k(n,q)$ third bad seeds have $|E_{x_0} \cup E_{x_1} \cup E_{x_2}| < d_k(\mathcal{C})$.

The expected size of the union of three independently sampled error sets is at most $n\Psi(\gamma)$, where $\Psi(\gamma) = 1-(1-\gamma)^3$. If $d_k(\mathcal{C}) > n\Psi(\gamma)$, Markov's inequality shows that at least a $1-n\Psi(\gamma)/d_k(\mathcal{C})$ fraction of ordered triples satisfy $|E_{x_0} \cup E_{x_1} \cup E_{x_2}| < d_k(\mathcal{C})$. For each fixed pair of distinct seeds x_0, x_1, we have just bounded the number of distinct third seeds satisfying this inequality by $\mathsf{T}_k(n,q)$. Comparing these lower and upper bounds, with an additional term for triples containing repeated seeds, gives

$$|\mathbb{K}|\epsilon_{\mathsf{MCA}}^{\mathcal{C}_{\mathbb{K}}}(\gamma) \le \frac{\mathsf{T}_k(n,q)+3}{1-n\Psi(\gamma)/d_k(\mathcal{C})} \ .$$

For fixed q, k and a fixed positive margin, this is polynomial in n, regardless of the list size; see Lemma 10.5.

Constructing the required generalized weights. The preceding bound reduces the task to constructing a code with a large list and a fixed generalized weight satisfying $d_k(\mathcal{C})/n > \Psi(\gamma)$ by a fixed positive margin. Put $H := \mathbb{F}_\ell^h$. A symmetric $(t+1)$-linear form is a map $T : H^{t+1} \to \mathbb{F}_\ell$ that is linear in each argument and unchanged by any permutation of its arguments. Denoting the dual of H by $H^\vee$, to each such form we associate the word

$$\Phi_T(x_1,\dots,x_t) := T(x_1,\dots,x_t,\cdot) \in H^\vee, \qquad (x_1,\dots,x_t) \in H^t \ .$$

Thus each coordinate is a linear functional on H, represented by a vector of h field elements. We construct an auxiliary code of dimension $\Theta(h^{t+1})$ consisting of such words. Each vector-valued coordinate is treated as one block, and relative distances count the fraction of blocks on which two words differ. The code has minimum relative block distance δ_t and a word with 2^{h-1} codewords at relative block distance β_t. Moreover, for any fixed $\eta > 0$, every k-dimensional subcode has relative block support at least $G_t - \eta$, for a sufficiently large fixed k independent of h. This last statement is exactly a lower bound on the k-th generalized weight, with whole blocks counted as coordinates. The useful relations between these three bounds are

$$\frac{\beta_t}{\delta_t} \longrightarrow \frac{1}{2} \ , \qquad \frac{G_t}{\delta_t} \longrightarrow \infty \qquad (t \to \infty) \ .$$

We next pass to a fixed extension $\mathbb{F}_q/\mathbb{F}_\ell$ and apply independently sampled linear functionals to the forms. Each functional is defined by sampling evaluation blocks and taking a random linear combination of their entries. The sampling parameters can be chosen to give minimum distance slightly above the target $\delta_{\mathcal{C}}$. For any small fixed $\varepsilon > 0$, the resulting code $\mathcal{C} \subseteq \mathbb{F}_q^n$ contains a set $\mathcal{L}$ of 2^{h-1} codewords and admits a word $z \in \mathbb{F}_q^n$ such that

$$\max_{\mathbf{c}\in\mathcal{L}} \frac{\Delta(z,\mathbf{c})}{n} \le 1-(1-\delta_{\mathcal{C}})^{\beta_t/\delta_t}+\varepsilon \ ,$$
$$\frac{d_k(\mathcal{C})}{n} \ge 1-(1-\delta_{\mathcal{C}})^{G_t/\delta_t}-\varepsilon \ .$$

As t grows, the two right-hand sides without their ε terms tend to $J_2(\delta_{\mathcal{C}})$ and one, respectively. Thus we first choose a fixed t large enough that the list radius lies below p and the generalized-weight bound lies above $\Psi(\gamma)$, with positive margins. We then choose the field size and tolerances so that these margins survive.

Only $\Theta(h^{t+1})$ linear functionals are needed to preserve the dimension and all these bounds simultaneously. Each functional gives one coordinate, so the scalar code has blocklength $n = \Theta(h^{t+1})$ and dimension

$\Theta(h^{t+1})$, giving a rate bounded away from zero. Since t is fixed, the list size $|\mathcal{L}| = 2^{h-1}$ grows faster than every polynomial in n. Appending a small fraction of zero coordinates makes the relative distance tend to exactly $\delta_{\mathcal{C}}$, while preserving the strict list-radius and generalized-weight inequalities. Finally, scalar extension to $\mathbb{K}$ preserves these properties, and the preceding MCA bound gives error $\mathsf{poly}(n)/|\mathbb{K}|$. The construction works in every prescribed characteristic; see Theorem 10.1. The alternative BCH construction in Appendix C gives the same separation in characteristic two.

3 Preliminaries

We write $\mathbb{N}$ for the set of positive integers including zero, and for $n \in \mathbb{N}$, $[n] := \{1, \ldots, n\}$. Codes are defined over finite fields. We use $\mathbb{F}$ for an arbitrary finite field and $\mathbb{F}_q$ for a field with q elements. Bold letters denote vectors or matrices. For $\mathbf{v} \in \mathbb{F}^n$ and $i \in [n]$, its i-th coordinate is $\mathbf{v}[i]$. For $T \subseteq [n]$, the restriction $\mathbf{v}|_T$ lists these coordinates in the order inherited from $[n]$. We write $\mathsf{Mat}(\mathbf{v}_0, \ldots, \mathbf{v}_{\ell-1})$ for the matrix whose rows are the indicated vectors. Probabilities and expectations indexed by a finite nonempty set refer to the uniform distribution on that set. Logarithms are natural unless a base is specified.

We write $V^\vee$ for the linear dual of a vector space V, and $A^\vee : W^\vee \to V^\vee$ for the dual of a linear map $A : V \to W$. For a chosen basis $e_1, \ldots, e_m$, the dual basis is denoted by $e_1^\vee, \ldots, e_m^\vee$.

3.1 Linear codes and list-decoding

A code of blocklength $n \geq 1$ over $\mathbb{F}$ is a nonempty subset $\mathcal{C} \subseteq \mathbb{F}^n$. Its elements are codewords, while arbitrary elements of $\mathbb{F}^n$ are words. The code is linear if it is an $\mathbb{F}$-linear subspace. Its dimension is $\dim_{\mathbb{F}} \mathcal{C}$ and its rate is $\dim_{\mathbb{F}} \mathcal{C}/n$.

The Hamming distance and relative Hamming distance between two words are

$$\Delta(\mathbf{u}, \mathbf{v}) := |\{i \in [n] : \mathbf{u}[i] \neq \mathbf{v}[i]\}| \ , \qquad \delta(\mathbf{u}, \mathbf{v}) := \frac{\Delta(\mathbf{u}, \mathbf{v})}{n} \ .$$

Their relative agreement is $\mathsf{agr}(\mathbf{u}, \mathbf{v}) := 1 - \delta(\mathbf{u}, \mathbf{v})$. We write

$$\Delta(\mathbf{u}, \mathcal{C}) := \min_{\mathbf{c} \in \mathcal{C}} \Delta(\mathbf{u}, \mathbf{c}) \qquad \text{and} \qquad \delta(\mathbf{u}, \mathcal{C}) := \frac{\Delta(\mathbf{u}, \mathcal{C})}{n} \ .$$

A word is p-close to $\mathcal{C}$ if its relative distance to $\mathcal{C}$ is at most p, and p-far otherwise. For a code with at least two codewords, its minimum distance and relative minimum distance are

$$\Delta_{\mathcal{C}} := \min_{\substack{\mathbf{c}, \mathbf{c}' \in \mathcal{C} \\ \mathbf{c} \neq \mathbf{c}'}} \Delta(\mathbf{c}, \mathbf{c}') \ , \qquad \delta_{\mathcal{C}} := \frac{\Delta_{\mathcal{C}}}{n} \ .$$

We also use $\Delta(\mathcal{C})$ and $\delta(\mathcal{C})$ for these two quantities. The support of a vector is $\mathsf{supp}(\mathbf{v}) := \{i : \mathbf{v}[i] \neq 0\}$ and its weight is $\mathsf{wt}(\mathbf{v}) := |\mathsf{supp}(\mathbf{v})|$. For a nonzero linear code, minimum distance equals the minimum weight of a nonzero codeword.

For $T \subseteq [n]$, the restricted code is

$$\mathcal{C}|_T := \{\mathbf{c}|_T : \mathbf{c} \in \mathcal{C}\} \ .$$

Thus $\mathbf{u}|_T \in \mathcal{C}|_T$ means that some codeword agrees with $\mathbf{u}$ on all of T. Equivalently, $\delta(\mathbf{u}, \mathcal{C}) \leq p$ if and only if this membership holds for some T of size at least $(1-p)n$. An information set of a linear code of dimension k is a set $I \subseteq [n]$ of size k for which restriction gives an isomorphism $\mathcal{C} \to \mathbb{F}^I$. Such a set exists by choosing k linearly independent columns of a matrix whose rows form a basis of $\mathcal{C}$.

Definition 3.1 (List-decoding)**.** *Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a code. For a relative radius $p \in [0, 1]$, its maximal list size is*

$$L_{\mathcal{C}}(p) := \max_{\mathbf{u} \in \mathbb{F}^n} |\{\mathbf{c} \in \mathcal{C} : \Delta(\mathbf{u}, \mathbf{c}) \leq pn\}| \ .$$

The code is (p, L)-list-decodable if $L_{\mathcal{C}}(p) \leq L$.

Unless explicitly stated otherwise, decoding and proximity radii are relative. Since Hamming distance is integral, a ball of relative radius p has absolute radius $\lfloor pn \rfloor$. Likewise, the condition $|T| \geq (1-p)n$ means $|T| \geq n - \lfloor pn \rfloor$. Assumptions that pn is integral in subsequent arguments are only an arithmetic convenience. The triangle inequality gives unique decoding whenever $2pn < \Delta_{\mathcal{C}}$: in this range, $L_{\mathcal{C}}(p) = 1$.

For an integer $h \geq 2$ and $\delta \in [0,1]$, put

$$J_h(\delta) := 1 - (1-\delta)^{1/h} \ .$$

The usual Johnson radius is $J_2(\delta_{\mathcal{C}}) = 1 - \sqrt{1-\delta_{\mathcal{C}}}$. The higher-order radii $J_h(\delta_{\mathcal{C}})$ arise in our MCA bounds; for the degree-d powers generator we take $h = d+2$ and call this the generalized Johnson radius.

Lemma 3.2 (Johnson bound)**.** *Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a code of relative minimum distance $\delta_{\mathcal{C}} > 0$. For $0 \leq p < J_2(\delta_{\mathcal{C}})$,*

$$L_{\mathcal{C}}(p) \leq \frac{\delta_{\mathcal{C}}}{(1-p)^2 - (1-\delta_{\mathcal{C}})} \ .$$

In particular, if $p \leq J_2(\delta_{\mathcal{C}}) - \varepsilon$ for some $\varepsilon > 0$, then $L_{\mathcal{C}}(p) \leq \varepsilon^{-2}$.

3.2 Generators and mutual correlated agreement

Definition 3.3 (Generator)**.** *A generator with output size $\ell \geq 1$ over $\mathbb{F}$ is a function $G : S \to \mathbb{F}^{\ell}$, where S is a finite nonempty seed set. Given words $\mathbf{u}_0, \ldots, \mathbf{u}_{\ell-1} \in \mathbb{F}^n$ and their matrix $\mathbf{U} := \mathsf{Mat}(\mathbf{u}_0, \ldots, \mathbf{u}_{\ell-1})$, its combined word at seed x is*

$$\mathbf{u}_x := G(x) \cdot \mathbf{U} = \sum_{j=0}^{\ell-1} G(x)_j \mathbf{u}_j \ .$$

The generators studied in this paper are the univariate powers generators

$$\begin{aligned} G_d &: S \to \mathbb{F}^{d+1} \\ G_d(x) &:= (1, x, \ldots, x^d) \ , \end{aligned}$$

where $d \geq 1$ and $\varnothing \neq S \subseteq \mathbb{F}$. Their combined words are evaluations of $\mathbf{u}(\mathtt{X}) := \sum_{j=0}^{d} \mathbf{u}_j \mathtt{X}^j$. The line generator is the case $d = 1$, with combined word $\mathbf{u}_0 + x\mathbf{u}_1$. Unless a seed set is specified, the line generator uses $S = \mathbb{F}$.

For the following definition, we also allow alphabets Σ that are finite-dimensional $\mathbb{F}$-vector spaces; an $\mathbb{F}$-linear code $\mathcal{C} \subseteq \Sigma^n$ is an $\mathbb{F}$-linear subspace. Hamming distance counts coordinates with unequal symbols in Σ; restriction and list size are defined as above with $\mathbb{F}^n$ replaced by Σ^n.

Definition 3.4 (Mutual correlated agreement [ACFY25])**.** *A generator $G\colon S \to \mathbb{F}^{\ell}$ has mutual correlated agreement (MCA) for an $\mathbb{F}$-linear code $\mathcal{C} \subseteq \Sigma^n$ with error $\epsilon_{\mathsf{MCA}}\colon [0,1] \to [0,1]$ if for every $\mathbf{u}_1, \ldots, \mathbf{u}_{\ell} \in \Sigma^n$ and $\gamma \in [0,1]$, defining the matrix $\mathbf{U} \coloneqq \mathsf{Mat}(\mathbf{u}_1, \ldots, \mathbf{u}_{\ell}) \in \Sigma^{\ell \times n}$, the following holds:*

$$\Pr_{\mathbf{x} \in S} \left[\begin{array}{l} \exists\, T \subseteq [n] \text{ s.t.} \\ \quad |T| \geq n \cdot (1-\gamma) \\ \quad \wedge\, (G(\mathbf{x}) \cdot \mathbf{U})|_T \in \mathcal{C}|_T \\ \quad \wedge\, \exists\, j \in [\ell], \mathbf{u}_j|_T \notin \mathcal{C}|_T \end{array} \right] \leq \epsilon_{\mathsf{MCA}}(\gamma) \ .$$

We denote the smallest such error by $\epsilon_{\mathsf{MCA}}^{\mathcal{C},G}(\gamma)$. We omit the generator, or both the code and the generator, when they are clear from context. In particular, $\epsilon_{\mathsf{MCA}}^{\mathcal{C}}(\gamma)$ denotes the MCA error of the line generator over $\mathbb{F}$ unless stated otherwise.

For fixed words, a seed is MCA-bad if it admits a set T witnessing the event in Definition 3.4. Thus, except for a $\epsilon_{\mathsf{MCA}}(\gamma)$ fraction of seeds, the required implication holds simultaneously for every sufficiently large T. Both the bad seeds and their witness sets may depend on the words. The minimum MCA error is the maximum bad-seed fraction over all word matrices. It is nondecreasing in γ, since increasing the radius only allows more witness sets.

3.3 Elementary bounds

We record two standard bounds used below.

Lemma 3.5 (DeMillo–Lipton–Schwartz–Zippel)**.** *Let $P \in \mathbb{F}[\mathtt{X}_1, \ldots, \mathtt{X}_a]$ be nonzero of total degree at most D, with $a \geq 1$. For a finite nonempty $S \subseteq \mathbb{F}$,*

$$\Pr_{(x_1,\ldots,x_a)\in S^a}[P(x_1, \ldots, x_a) = 0] \leq \frac{D}{|S|} \ .$$

In particular, a nonzero univariate polynomial of degree at most D has at most D roots, and if $|\mathbb{F}| > D$, then P is nonzero at some point of $\mathbb{F}^a$.

Lemma 3.6 (Hoeffding's inequality)**.** *Let $N \geq 1$ and let $Y_1, \ldots, Y_N$ be independent random variables in $[0, 1]$. Put $\overline{Y} := N^{-1} \sum_{i=1}^{N} Y_i$. For every $\eta > 0$,*

$$\Pr[\overline{Y} - \mathbb{E}[\overline{Y}] \geq \eta] \leq e^{-2N\eta^2} \ ,$$
$$\Pr[\overline{Y} - \mathbb{E}[\overline{Y}] \leq -\eta] \leq e^{-2N\eta^2} \ .$$

Consequently, $\Pr[|\overline{Y} - \mathbb{E}[\overline{Y}]| \geq \eta] \leq 2e^{-2N\eta^2}$.

4 Bad seeds, curves, and root counting

Fix a finite-dimensional $\mathbb{F}$-vector space Σ and an $\mathbb{F}$-linear code $\mathcal{C} \subseteq \Sigma^n$ of minimum Hamming distance $\Delta_{\mathcal{C}}$ and relative distance $\delta_{\mathcal{C}} := \Delta_{\mathcal{C}}/n$. Fix $d \geq 1$ and put $\ell := d+1$ and $h := d+2$. We use $\gamma \in [0,1]$ for the proximity radius and initially assume that γn is an integer. For an arbitrary real radius, all witness conditions depend only on $\lfloor \gamma n \rfloor$, so the results below apply with γ replaced by $\lfloor \gamma n \rfloor / n$:

$$\gamma n \in \{0, \ldots, n\} \ . \tag{2}$$

We fix words

$$\begin{aligned} &\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_d \in \Sigma^n \ , \\ &\mathbf{U} := \mathsf{Mat}(\mathbf{u}_0, \ldots, \mathbf{u}_d) \in \Sigma^{\ell \times n} \ , \end{aligned}$$

so that $G_d(x) \cdot \mathbf{U} = \mathbf{u}(x)$ where

$$\mathbf{u}(\mathtt{X}) := \sum_{j=0}^{d} \mathbf{u}_j \mathtt{X}^j$$

is a vector of n univariate polynomials with coefficients in Σ, each of degree at most d; we write $\mathbf{u}_i(\mathtt{X})$ for its i-th coordinate. All polynomial identities below are formal coefficientwise equalities.

Definition 4.1 (Bad seeds)**.** *The set of* MCA-bad seeds *for* $\mathbf{U}$ *and* γ *is*

$$\mathsf{B} = \mathsf{B}_{\mathcal{C}}(\mathbf{U}; \gamma) := \left\{ x \in S \ : \ \begin{array}{l} \exists\, T \subseteq [n] \text{ with } |T| \geq n(1-\gamma) \text{ such that} \\ \mathbf{u}(x)|_T \in \mathcal{C}|_T \text{ and } \mathbf{u}_j|_T \notin \mathcal{C}|_T \text{ for some } j \in \{0, \ldots, d\} \end{array} \right\} \ .$$

By Definition 4.1, bounding $|\mathsf{B}|$ uniformly over $\mathbf{U}$ and dividing by $|S|$ bounds $\epsilon_{\mathsf{MCA}}(\gamma)$.

Lemma 4.2 (Decoding and full agreement sets)**.** *Assume* $2\gamma n < \Delta_{\mathcal{C}}$*. For every* $x \in \mathsf{B}$ *there is a unique codeword* $\mathbf{c}_x \in \mathcal{C}$ *with* $\Delta(\mathbf{u}(x), \mathbf{c}_x) \leq \gamma n$*. Moreover, letting*

$$A(x) := \{ i \in [n] \ : \ \mathbf{u}_i(x) = \mathbf{c}_x[i] \} \ ,$$

the set $A(x)$ *is itself a witness for the badness of* x*:* $|A(x)| \geq n(1-\gamma)$*,* $\mathbf{u}(x)|_{A(x)} \in \mathcal{C}|_{A(x)}$*, and* $\mathbf{u}_j|_{A(x)} \notin \mathcal{C}|_{A(x)}$ *for some* j*.*

Proof. Let T be a witness for $x \in \mathsf{B}$ and let $\mathbf{c} \in \mathcal{C}$ satisfy $\mathbf{u}(x)|_T = \mathbf{c}|_T$. Then $\Delta(\mathbf{u}(x), \mathbf{c}) \leq n - |T| \leq \gamma n$, and since $2\gamma n < \Delta_{\mathcal{C}}$ such a codeword is unique; set $\mathbf{c}_x := \mathbf{c}$. By construction $T \subseteq A(x)$, so $|A(x)| \geq |T| \geq n(1-\gamma)$. Finally, if $\mathbf{u}_j|_T \notin \mathcal{C}|_T$ then a fortiori $\mathbf{u}_j|_{A(x)} \notin \mathcal{C}|_{A(x)}$, because a codeword agreeing with $\mathbf{u}_j$ on $A(x) \supseteq T$ would in particular agree with it on T. □

Lemma 4.3 (The case $\gamma = 0$)**.** *If* $\gamma = 0$ *then* $|\mathsf{B}| \leq d$*.*

Proof. If every $\mathbf{u}_j$ lies in $\mathcal{C}$ then $\mathsf{B} = \emptyset$. Otherwise the polynomial $\mathbf{u}(\mathtt{X}) + \mathcal{C}$ with coefficients in the quotient space $\Sigma^n/\mathcal{C}$ is nonzero, so some $\mathbb{F}$-linear functional φ on $\Sigma^n/\mathcal{C}$ takes it to a nonzero polynomial $\varphi(\mathbf{u}(\mathtt{X}) + \mathcal{C}) \in \mathbb{F}[\mathtt{X}]$ of degree at most d. Every $x \in \mathsf{B}$ satisfies $\mathbf{u}(x) \in \mathcal{C}$ and is therefore a root of that polynomial. □

From now on we assume $\gamma n \geq 1$, $2\gamma n < \Delta_{\mathcal{C}}$ and $\gamma < J_h(\delta_{\mathcal{C}})$. We collect below several useful facts from the literature on proximity gaps and mutual correlated agreement [GKL24; Hab25; BCGM26; GCXK26; GG26; GGSW26; Zei26], with proofs in our notation.

Definition 4.4 (Curves)**.** *A* codeword polynomial*, or* curve*, is a formal polynomial*

$$P(\mathtt{X}) = \sum_{j=0}^{d} P_j \mathtt{X}^j \ ,$$
$$P_j \in \mathcal{C} \ .$$

Its domain *is* $T_P := \{i \in [n] : \mathbf{u}_i(\mathtt{X}) \equiv P_i(\mathtt{X})\}$*, and we write* $z_P := |T_P|/n$ *and* $s_P := n - |T_P|$*. We say that a seed* $x \in \mathsf{B}$ lies on P *if* $\mathbf{c}_x = P(x)$*.*

Lemma 4.5 (Root counting on a curve)**.** *Let* P *be a curve and let* $\mathsf{B}' \subseteq \mathsf{B}$*. Then*

$$|\{x \in \mathsf{B}' : \mathbf{c}_x = P(x)\}| \ \le \ d \cdot (\gamma n + 1) \ .$$

If moreover $z_P < (1-\gamma)$*, then*

$$|\{x \in \mathsf{B}' : \mathbf{c}_x = P(x)\}| \ \le \ d \cdot \frac{1 - z_P}{1 - \gamma - z_P} \ = \ d + \frac{d\gamma}{1 - \gamma - z_P} \ .$$

Proof. Write $s := s_P$ and let r be the number of counted seeds. If $s = 0$ then $\mathbf{u}_j = P_j \in \mathcal{C}$ for every j and $\mathsf{B} = \emptyset$. So assume $s \ge 1$; the s coordinate polynomials $\mathbf{u}_i(\mathtt{X}) - P_i(\mathtt{X})$ with $i \notin T_P$ are nonzero of degree at most d. Each has at most d roots: applying an $\mathbb{F}$-linear functional on Σ that is nonzero on one coefficient gives a nonzero scalar polynomial of degree at most d.

Let x be a counted seed. Since $\Delta(\mathbf{u}(x), \mathbf{c}_x) \le \gamma n$ and $\mathbf{c}_x = P(x)$, at least $s - \gamma n$ of these s polynomials vanish at x. At least one of them vanishes at x even when $s \le \gamma n$: otherwise $A(x) = T_P$, and then $\mathbf{u}_j|_{A(x)} = P_j|_{A(x)} \in \mathcal{C}|_{A(x)}$ for every j, contradicting Lemma 4.2. Counting root incidences therefore gives

$$r \cdot \max\{1, s - \gamma n\} \ \le \ d\, s \ .$$

If $s \le \gamma n + 1$ this yields $r \le ds \le d(\gamma n + 1)$, and if $s \ge \gamma n + 1$ it yields $r \le d\frac{s}{s - \gamma n} = d\left(1 + \frac{\gamma n}{s - \gamma n}\right) \le d(\gamma n + 1)$, which is the first bound. For the second, $z_P < (1-\gamma)$ means $s - \gamma n = n(1 - \gamma - z_P) > 0$, so $r \le d\frac{s}{s-\gamma n} = d\frac{1-z_P}{1-\gamma-z_P}$, and $\frac{1-z_P}{1-\gamma-z_P} = 1 + \frac{\gamma}{1-\gamma-z_P}$. □

Lemma 4.6 (Interpolation)**.** *Let* $x_1, \ldots, x_\ell \in \mathsf{B}$ *be distinct. There is a unique curve* P *with* $P(x_m) = \mathbf{c}_{x_m}$ *for all* $m \in [\ell]$*, and*

$$A(x_1) \cap \cdots \cap A(x_\ell) = T_P \ .$$

Proof. Lagrange interpolation over $\mathbb{F}$ through $\ell = d + 1$ distinct points produces a unique polynomial of degree at most d whose coefficients are $\mathbb{F}$-linear combinations of $\mathbf{c}_{x_1}, \ldots, \mathbf{c}_{x_\ell}$, hence lie in $\mathcal{C}$ by linearity of $\mathcal{C}$. If $i \in T_P$ then $\mathbf{u}_i(\mathtt{X}) \equiv P_i(\mathtt{X})$ and in particular $\mathbf{u}_i(x_m) = P_i(x_m) = \mathbf{c}_{x_m}[i]$ for all m. Conversely, if i lies in all the $A(x_m)$ then $\mathbf{u}_i(\mathtt{X}) - P_i(\mathtt{X})$ is a polynomial of degree at most d with $d + 1$ distinct roots, hence identically zero. □

Definition 4.7 (Coherent tuples)**.** *A tuple of* $h = \ell + 1$ *distinct seeds of* B *is* coherent *if all of its decoded points lie on a single curve, and* noncoherent *otherwise.*

Remark 4.8 (Relation to curve-decodability)**.** Coherence describes individual configurations in the curve-decodability framework of [GG26], developed further in [GGSW26]. In our notation, (d, γ, a, b) curve-decodability requires that whenever a choice of codeword witnesses gives at least a evaluations within relative

distance γ of an ambient degree-d curve, at least b of those parameterized witnesses lie on one codeword curve. A coherent tuple is such a configuration with $b = d + 2$, the first size for which the condition is not automatic by interpolation. For lines, the corresponding code property is also called collinearity of proximates [Hab24].

There is an exact correspondence with the curve-free subspaces of [GGSW26]. Fix $A \subseteq \mathsf{B}$ with $|A| \geq d + 2$ and choose an $\mathbb{F}$-basis of Σ. Let $W_A \subseteq \mathbb{F}^A$ be the row space of the scalar matrix whose column indexed by x is the basis-coordinate representation of $\mathbf{c}_x \in \Sigma^n$. For $B \subseteq A$ of size $d + 2$, let $\mathsf{RS}_{\leq d}(B)$ be the space of evaluations on B of scalar polynomials of degree at most d. Then

$$B \text{ is coherent} \quad \Longleftrightarrow \quad W_A|_B \subseteq \mathsf{RS}_{\leq d}(B) \ .$$

Indeed, the inclusion says that the decoded points interpolate coordinatewise to a degree-d curve; its coefficients lie in $\mathcal{C}$ by linearity. Thus W_A is $(d, A, d + 2)$ curve-free precisely when every h-tuple in A is noncoherent.

Our argument permits coherent tuples and bounds their frequency. In particular, removing large curves need not make the associated row space curve-free, since small curves may still contain h decoded points.

Lemma 4.9 (Noncoherent tuples)**.** *The common agreement set of a noncoherent h-tuple of seeds has size at most $n - \Delta_{\mathcal{C}}$.*

Proof. Let $x_1, \ldots, x_h$ be the tuple, let P be the curve interpolating the first ℓ decoded points (Lemma 4.6), and write $y := x_h$. Noncoherence means that $\mathbf{c}_y - P(y) \in \mathcal{C}$ is nonzero. Let i be in the common agreement set. By Lemma 4.6, $i \in T_P$, so $\mathbf{u}_i(\mathtt{X}) \equiv P_i(\mathtt{X})$ and in particular $\mathbf{u}_i(y) = P_i(y)$; since also $i \in A(y)$, we get $\mathbf{c}_y[i] = P_i(y)$. Thus the nonzero codeword $\mathbf{c}_y - P(y)$ vanishes on the common agreement set, whose size is therefore at most $n - \Delta_{\mathcal{C}}$. □

Lemma 4.10 (A curve with a point outside it)**.** *Let P be a curve and suppose that some seed of B does not lie on P. Then*

$$s_P \geq \Delta_{\mathcal{C}} - \gamma n \ ,$$
$$\textit{equivalently, } z_P \leq 1 - \delta_{\mathcal{C}} + \gamma \ .$$

Proof. Let $y \in \mathsf{B}$ with $\mathbf{c}_y \neq P(y)$, so that $\mathbf{c}_y - P(y)$ is a nonzero codeword. If $i \in T_P$ and $i \in A(y)$ then, as in the proof of Lemma 4.9, $\mathbf{c}_y[i] = P_i(y)$. Hence the support of $\mathbf{c}_y - P(y)$ is contained in the union of $[n] \setminus T_P$ and $[n] \setminus A(y)$, so $\Delta_{\mathcal{C}} \leq s_P + \gamma n$. □

We finally record a Corrádi-type set-intersection bound [Juk11] that controls the curves with a large domain. We have $\sqrt{1 - \delta_{\mathcal{C}}} \leq (1 - \delta_{\mathcal{C}})^{1/3} \leq (1 - \delta_{\mathcal{C}})^{1/h}$, the first inequality because $1 - \delta_{\mathcal{C}} \leq 1$ and the second because $h \geq 3$. Since $\gamma < J_h(\delta_{\mathcal{C}})$ we have $(1 - \gamma) > (1 - \delta_{\mathcal{C}})^{1/h} \geq \sqrt{1 - \delta_{\mathcal{C}}}$, and since $2\gamma n < \Delta_{\mathcal{C}}$ we have $\gamma < \delta_{\mathcal{C}}/2$, i.e.

$$(1 - \gamma) \ > \ \max\left\{(1 - \delta_{\mathcal{C}})^{1/3}, \ \tfrac{1}{2}\left(1 + (1 - \delta_{\mathcal{C}})\right)\right\} \ . \tag{3}$$

Lemma 4.11 (Few large curves, Corrádi-type bound)**.** *Let $\tau \in (\sqrt{1 - \delta_{\mathcal{C}}}, 1)$. Then the number of curves P with $z_P > \tau$ is at most*

$$\frac{\tau - (1 - \delta_{\mathcal{C}})}{\tau^2 - (1 - \delta_{\mathcal{C}})} \ .$$

In particular, if $\tau = \frac{1 - \gamma + \sqrt{1 - \delta_{\mathcal{C}}}}{2}$ then there are at most 4 such curves.

Proof. If $P \neq P'$ are curves then some coefficient of $P - P'$ is a nonzero codeword, and that coefficient vanishes on $T_P \cap T_{P'}$; hence

$$|T_P \cap T_{P'}| \leq n - \Delta_{\mathcal{C}} = (1 - \delta_{\mathcal{C}})n \ .$$

Let $P_1, \ldots, P_L$ be distinct curves with $z_{P_m} > \tau$ and let $\bar{z} := \frac{1}{L}\sum_m z_{P_m} \geq \tau$. For $i \in [n]$ let $\nu_i := |\{m : i \in T_{P_m}\}|$. Then $\sum_i \nu_i = L\bar{z}n$ and $\sum_i \nu_i(\nu_i - 1) \leq L(L-1)(1-\delta_{\mathcal{C}})n$, while Cauchy–Schwarz gives $\sum_i \nu_i^2 \geq L^2\bar{z}^2 n$. Combining,

$$\begin{aligned} &L\bar{z}^2 - \bar{z} \leq (L-1)(1-\delta_{\mathcal{C}}) \ , \\ &\text{i.e., } L\left(\bar{z}^2 - (1-\delta_{\mathcal{C}})\right) \leq \bar{z} - (1-\delta_{\mathcal{C}}) \ . \end{aligned}$$

The map $z \mapsto \frac{z-(1-\delta_{\mathcal{C}})}{z^2-(1-\delta_{\mathcal{C}})}$ is decreasing on $(\sqrt{1-\delta_{\mathcal{C}}}, 1)$, so $L \leq \frac{\tau-(1-\delta_{\mathcal{C}})}{\tau^2-(1-\delta_{\mathcal{C}})}$. The last part is proven in Lemma A.1. □

The proofs of Lemmas 4.5, 4.6, 4.9 and 4.10 also apply above the unique decoding radius: choose one witness codeword for each bad seed and let $A(x)$ be its full agreement set. This set remains a badness witness because it contains the original witness set. The bound of four large curves still requires (3).

For the rest of this section, assume instead that an $\mathbb{F}$-linear code $\mathcal{C} \subseteq \Sigma^n$ is (p, L)-list-decodable, where $L \geq 1$ is an integer and $|\mathbb{F}| > L$. Fix $\mathbf{v}, \mathbf{w} \in \Sigma^n$ and $0 \leq \gamma \leq p$, put $\mathsf{B} := \mathsf{B}_{\mathcal{C}}(\mathbf{v}, \mathbf{w}; \gamma)$, and let $\mathbf{u}(\mathtt{X}) := \mathbf{v} + \mathtt{X}\mathbf{w}$. For a codeword line $Q(\mathtt{X}) = \mathbf{q}_0 + \mathtt{X}\mathbf{q}_1$ with $\mathbf{q}_0, \mathbf{q}_1 \in \mathcal{C}$, we again let

$$T_Q = \{i \in [n] : \mathbf{u}_i(\mathtt{X}) \equiv Q_i(\mathtt{X})\} \ ,$$

For each bad seed x, choose a witness codeword $\mathbf{c}_x$ and let $A(x)$ be its full agreement set with $\mathbf{u}(x)$. As above, this set is still a witness for MCA-badness. We say that x lies on Q if $\mathbf{c}_x = Q(x)$, and write m_Q for the number of such seeds. We call a codeword line Q large if $|T_Q| \geq n(1-p)$. Then both of the following results hold: (i) $m_Q \leq \lfloor n\gamma \rfloor + 1$ for every large codeword line Q, and (ii) the number of large codeword lines is at most $\frac{L(|\mathbb{F}|-1)}{|\mathbb{F}|-L}$. We record their proofs for completeness.

Lemma 4.12 (Root counting on a codeword line)**.** *For every codeword line Q, let $s_Q := n - |T_Q|$. Then,*

$$\begin{aligned} &m_Q \cdot \max\{1, s_Q - \gamma n\} \ \leq \ s_Q \ , \\ &\textit{and in particular, } m_Q \ \leq \ \gamma n + 1 \ . \end{aligned} \tag{4}$$

Proof. If $s_Q = 0$ then $\mathbf{v} = \mathbf{q}_0 \in \mathcal{C}$ and $\mathbf{w} = \mathbf{q}_1 \in \mathcal{C}$, so $\mathsf{B} = \varnothing$. Assume $s_Q > 0$, so that the s_Q coordinate polynomials $\mathbf{u}_i - Q_i$ with $i \notin T_Q$ are nonzero affine polynomials with coefficients in Σ, each with at most one root.

Let x be a seed lying on Q. For $i \in A(x)$ we have $\mathbf{u}_i(x) = \mathbf{c}_x[i] = Q_i(x)$, so $\mathbf{u}_i - Q_i$ vanishes at x; hence x is a root of at least $|A(x) \setminus T_Q| \geq s_Q - \gamma n$ of the polynomials above. It is also a root of at least one of them even when $s_Q \leq \gamma n$: otherwise $A(x) \subseteq T_Q$, and since $T_Q \subseteq A(x)$ always, $A(x) = T_Q$; but on T_Q the words $\mathbf{v}$ and $\mathbf{w}$ agree with $\mathbf{q}_0$ and $\mathbf{q}_1$ respectively, contradicting the badness of x. Counting root incidences gives the first inequality in (4).

For the second, if $s_Q \leq \gamma n + 1$ then $m_Q \leq s_Q \leq \gamma n + 1$, and otherwise $m_Q \leq s_Q/(s_Q - \gamma n) = 1 + \gamma n/(s_Q - \gamma n) \leq \gamma n + 1$. □

Lemma 4.13 (Few large lines)**.** *The number of large codeword lines is at most $A_L := \frac{L(|\mathbb{F}|-1)}{|\mathbb{F}|-L}$.*

Proof. Let $Q^{(1)}, \ldots, Q^{(N)}$ be distinct large codeword lines. For every $x \in \mathbb{F}$ and every m, the codeword $Q^{(m)}(x)$ agrees with $\mathbf{u}(x)$ on $T_{Q^{(m)}}$, hence $\Delta(Q^{(m)}(x), \mathbf{u}(x)) \leq pn$. By list-decodability the N codewords $Q^{(1)}(x), \ldots, Q^{(N)}(x)$ take at most L distinct values, so by Cauchy–Schwarz there are at least $N^2/L - N$ ordered pairs (m, m') with $m \neq m'$ and $Q^{(m)}(x) = Q^{(m')}(x)$. On the other hand two distinct affine vector polynomials agree at no more than one seed, so summing over $x \in \mathbb{F}$,

$$|\mathbb{F}| \left(\frac{N^2}{L} - N \right) \;\leq\; N(N-1) \;\; .$$

Dividing by N and rearranging gives $N(|\mathbb{F}| - L) \leq L(|\mathbb{F}| - 1)$. □

5 Agreement moments

The lower bounds in this section depend only on the sizes and intersections of a family of sets. We first state them in this general setting, then apply them to agreement sets of bad seeds.

5.1 Moments of a family of sets

Let $\mathcal{S}$ be a finite index set of size $M \geq 1$, and let $A(x) \subseteq [n]$ for each $x \in \mathcal{S}$. For a coordinate $i \in [n]$, put

$$A_i := |\{x \in \mathcal{S} : i \in A(x)\}| \ .$$

Fix $0 \leq \gamma < 1$ and assume that

$$\frac{1}{n} \sum_{i \in [n]} A_i = \frac{1}{n} \sum_{x \in \mathcal{S}} |A(x)| \ \geq \ M(1-\gamma) \ . \tag{5}$$

In particular, this holds if every set has size at least $(1-\gamma)n$, but only the average size is needed below.

Definition 5.1 (Agreement moments)**.** *For an integer $1 \leq j \leq M$, the j-th agreement moment of the family $(A(x))_{x \in \mathcal{S}}$ is*

$$\mu_j := \frac{1}{n} \sum_{i \in [n]} \frac{(A_i)_j}{(M)_j} \ ,$$
$$(y)_j := y(y-1)\cdots(y-j+1) \ ,$$

with the convention that $(A_i)_j = 0$ when $A_i < j$ is a nonnegative integer. Equivalently,

$$\mu_j = \mathbb{E}\left[\frac{|A(X_1) \cap \cdots \cap A(X_j)|}{n}\right] \ ,$$

where $X_1, \ldots, X_j$ are sampled uniformly without replacement from $\mathcal{S}$.

The cubic agreement-moment argument with independently sampled seeds appears in [BGKS20]. Here we use factorial moments to account for sampling distinct indices without replacement.

Lemma 5.2 (Product bound)**.** *If $1 \leq j \leq M$ and $M(1-\gamma) \geq j$, then*

$$\mu_j \ \geq \ \frac{(M(1-\gamma))_j}{(M)_j} \ .$$

Proof. On the nonnegative integers the sequence $m \mapsto (m)_j = j!\binom{m}{j}$ is nondecreasing and discretely convex, so its piecewise-linear interpolation f is nondecreasing and convex on $[0, \infty)$. By Jensen's inequality and (5),

$$\frac{1}{n} \sum_i (A_i)_j = \frac{1}{n} \sum_i f(A_i) \ \geq \ f\left(\frac{1}{n} \sum_i A_i\right) \ \geq \ f(M(1-\gamma)) \ .$$

On $[j-1, \infty)$ the polynomial $y \mapsto (y)_j$ is convex: writing $y = w + j - 1$ gives $w(w+1)\cdots(w+j-1)$, which has nonnegative coefficients in $w \geq 0$. A convex function lies below its linear interpolation between consecutive integers, so $f(M(1-\gamma)) \geq (M(1-\gamma))_j$. Dividing by $(M)_j$ concludes. □

Corollary 5.3 (A lower bound on the moment root)**.** *Assume $j \geq 1$, $M \geq 2j$ and $1-\gamma > 1/2$. Then*

$$\mu_j^{1/j} \;\geq\; 1-\gamma-\frac{(j-1)\gamma}{M-j+1} \;\; .$$

Proof. We have $M(1-\gamma) > j$, so Lemma 5.2 gives

$$\begin{aligned}\mu_j &\geq \prod_{m=0}^{j-1} \frac{M(1-\gamma)-m}{M-m} \\ &= \prod_{m=0}^{j-1} \left(1-\gamma-\frac{m\gamma}{M-m}\right) \;\geq\; \left(1-\gamma-\frac{(j-1)\gamma}{M-j+1}\right)^j \;\; .\end{aligned}$$

Here $m \mapsto m\gamma/(M-m)$ is nondecreasing. The final base is positive, since $(j-1)\gamma/(M-j+1) < 1/2 < 1-\gamma$, so taking j-th roots proves the claim. □

Lemma 5.4 (Adjacent moments)**.** *For $1 \leq j < M$, set $v := \mu_j^{1/j}$. Then*

$$\mu_{j+1} \;\geq\; \frac{Mv^{j+1}-jv^j}{M-j} \;\; .$$

Proof. Put $w_i := (A_i)_j \,/\, (M)_j \in [0,1]$. If $A_i \geq j$, each factor $(A_i-m)/(M-m)$, for $0 \leq m < j$, is at most A_i/M. If $A_i < j$, then $w_i = 0$. Thus in both cases $w_i \leq (A_i/M)^j$, or equivalently $A_i \geq M w_i^{1/j}$. Using the falling-factorial identities,

$$\begin{aligned}\mu_{j+1} &= \frac{1}{M-j}\left(\frac{1}{n}\sum_i w_i A_i - j\mu_j\right) \\ &\geq \frac{1}{M-j}\left(M\cdot\frac{1}{n}\sum_i w_i^{1+1/j} - j\mu_j\right) \;\; .\end{aligned}$$

By convexity of $y \mapsto y^{1+1/j}$ and $n^{-1}\sum_i w_i = \mu_j$, the average in the last line is at least $\mu_j^{1+1/j} = v^{j+1}$. □

We next consider coordinates belonging to one sampled set and to none of several others. For these estimates, assume $M \geq 2$ and strengthen (5) to the equality

$$\frac{1}{n}\sum_i A_i = M(1-\gamma) \;\; . \tag{6}$$

Again, only the average size is prescribed; the individual sets need not have equal size.

Definition 5.5 (Agreement-disagreement moments)**.** *For an integer $k \geq 1$, sample $\mathcal{X}$ uniformly from $\mathcal{S}$ and, conditionally on $\mathcal{X}$, sample $Y_1, \ldots, Y_k$ independently and uniformly from $\mathcal{S} \setminus \{\mathcal{X}\}$. The k-th agreement-disagreement moment is*

$$\nu_k := \mathbb{E}\left[\frac{|A(\mathcal{X}) \setminus \bigcup_{j\in[k]} A(Y_j)|}{n}\right] = \frac{1}{n}\sum_{i=1}^n \frac{A_i}{M}\left(\frac{M-A_i}{M-1}\right)^k \;\; . \tag{7}$$

Each Y_j differs from $\mathcal{X}$, but the Y_j may repeat. The moment therefore measures agreement with one index and disagreement with k independently sampled other indices.

To see the coordinate formula, the probability that $i \in A(\mathcal{X})$ is A_i/M. Conditional on this event, exactly $M - A_i$ of the other $M - 1$ sets omit i, and the k samples are independent.

Lemma 5.6 (Agreement and agreement-disagreement estimates)**.** *Under* (6)*, for every integer* $k \geq 1$*,*

$$\begin{aligned} \mu_2 &\geq (1-\gamma)^2 - \frac{(1-\gamma)\gamma}{M-1} \ , \\ \nu_k &\geq (1-\gamma)\left(1 - \frac{\mu_2}{1-\gamma}\right)^k \ , \\ \mu_2 + \frac{\gamma + \nu_k}{k} &\geq F_k(\gamma) - \frac{(1-\gamma)\gamma}{M-1} \ , \end{aligned} \tag{8}$$

where

$$F_k(\gamma) := (1-\gamma)^2 + \frac{\gamma + (1-\gamma)\gamma^k}{k} \ .$$

Proof. For the first inequality, (6) and Cauchy–Schwarz give $n^{-1}\sum_i A_i^2 \geq M^2(1-\gamma)^2$. Hence

$$\mu_2 = \frac{n^{-1}\sum_i A_i^2 - M(1-\gamma)}{M(M-1)} \geq (1-\gamma)^2 - \frac{(1-\gamma)\gamma}{M-1} \ .$$

For the second, restrict to coordinates with $A_i > 0$ and put

$$\omega_i := \frac{A_i}{M(1-\gamma)n}, \qquad \theta_i := \frac{A_i - 1}{M-1} \ .$$

The equality in (6) ensures that the weights ω_i sum to one. Moreover,

$$\begin{aligned} \sum_{i:\, A_i > 0} \omega_i \theta_i &= \frac{\mu_2}{1-\gamma} \ , \\ \nu_k &= (1-\gamma) \sum_{i:\, A_i > 0} \omega_i (1-\theta_i)^k \ . \end{aligned}$$

Since $0 \leq \theta_i \leq 1$, Jensen's inequality for $z \mapsto (1-z)^k$ on $[0,1]$ gives the desired lower bound. It also follows that $0 \leq \mu_2 \leq 1 - \gamma$.

For the third inequality, define

$$\psi(z) := z + \frac{\gamma}{k} + \frac{1-\gamma}{k}\left(1 - \frac{z}{1-\gamma}\right)^k, \qquad 0 \leq z \leq 1-\gamma \ .$$

The second inequality gives $\mu_2 + (\gamma + \nu_k)/k \geq \psi(\mu_2)$. Also,

$$\psi'(z) = 1 - \left(1 - \frac{z}{1-\gamma}\right)^{k-1} \in [0,1] \ ,$$

where for $k = 1$ the derivative is identically zero. Thus ψ is nondecreasing and 1-Lipschitz. The first inequality now implies

$$\psi(\mu_2) \geq \psi((1-\gamma)^2) - \frac{(1-\gamma)\gamma}{M-1} \ .$$

Finally, $\psi((1-\gamma)^2) = F_k(\gamma)$. □

5.2 Agreement sets of bad seeds

Fix a finite-dimensional $\mathbb{F}$-vector space Σ, an $\mathbb{F}$-linear code $\mathcal{C} \subseteq \Sigma^n$, words $\mathbf{U} \in \Sigma^{\ell \times n}$ and a set $\mathsf{B}' \subseteq \mathsf{B}(\mathbf{U}; \gamma)$ of $M \geq 1$ bad seeds. Assume that γn is an integer. For each $x \in \mathsf{B}'$, choose a codeword $\mathbf{c}_x$ supplied by a badness witness, and let $A(x)$ be its full agreement set with $\mathbf{u}(x)$. Then $|A(x)| \geq (1-\gamma)n$. These choices do not require unique decoding; below the unique decoding radius they coincide with those of Lemma 4.2.

Apply Section 5.1 with $\mathcal{S} = \mathsf{B}'$ and $A(x) = A(x)$. Double counting gives (5), so the product and adjacent-moment bounds apply. We retain the notation μ_j for the resulting moments. In particular, for a degree-d generator, setting $j = \ell = d+1$ gives, when $M \geq 2\ell$ and $1 - \gamma > 1/2$,

$$v := \mu_\ell^{1/\ell} \;\geq\; 1 - \gamma - \theta, \qquad \theta := \frac{d\gamma}{M - d} \;,$$

and, when $M > \ell$,

$$\mu_{\ell+1} \;\geq\; \frac{Mv^{\ell+1} - \ell v^\ell}{M - \ell} \;.$$

These are the lower bounds used in Section 7; the properties of the code enter there through an upper bound on the same intersections.

For the agreement-disagreement estimates, we need the exact average in (6). Choose subsets

$$\widehat{A}(x) \subseteq A(x), \qquad |\widehat{A}(x)| = (1-\gamma)n \;,$$

and write $\alpha_i := |\{x \in \mathsf{B}' : i \in \widehat{A}(x)\}|$. Then

$$\sum_i \alpha_i = M(1-\gamma)n \;.$$

For $M \geq 2$, the pairwise agreement moment of this trimmed family is

$$\hat{\mu}_2 := \frac{1}{n} \sum_{i=1}^{n} \frac{(\alpha_i)_2}{(M)_2} = \mathbb{E}_{\substack{x,y \in \mathsf{B}' \\ x \neq y}} \left[\frac{|\widehat{A}(x) \cap \widehat{A}(y)|}{n} \right] \;. \tag{9}$$

The expectation is over an ordered pair sampled uniformly without replacement. The notation $\hat{\mu}_2$ distinguishes this quantity from μ_2 for the full agreement sets. In the next subsection, ν_k denotes the agreement-disagreement moment of the trimmed family, obtained by substituting α_i for A_i in (7). Thus Lemma 5.6 applies with μ_2 replaced by $\hat{\mu}_2$. The trimmed sets are used only for counting; we keep the full agreement sets as witnesses for badness.

5.3 List-decodability and the completion score

Suppose now that $\mathcal{C} \subseteq \mathbb{F}^n$ is linear and (p, L)-list-decodable, where $L \geq 1$ is an integer, $pn \in \{0, \ldots, n-1\}$ and $0 \leq \gamma \leq p$. For the line generator, fix $\mathbf{v}, \mathbf{w} \in \mathbb{F}^n$ and a set $\mathsf{B}' \subseteq \mathsf{B}(\mathbf{v}, \mathbf{w}; \gamma)$ of $M \geq 2$ bad seeds, together with the codeword witnesses and trimmed sets from Section 5.2.

Fix $x \in \mathsf{B}'$ and neighbors $y_1, \ldots, y_{L+1} \in \mathsf{B}' \setminus \{x\}$, allowing repetitions. The slope codewords

$$d_{x,y_j} := \frac{\mathbf{c}_{y_j} - \mathbf{c}_x}{y_j - x} \in \mathcal{C}$$

agree with $\mathbf{w}$ on $\widehat{A}(x) \cap \widehat{A}(y_j)$: this follows by subtracting the two witness equalities on that intersection. To put these codewords in one decoding ball, we retain these agreements and use the coordinates outside their union to supply additional ones. This leads to the following score.

Definition 5.7 (Completion score)**.** *For $x, y_1, \dots, y_{L+1}$ as above, put*

$$\begin{aligned} \ell_j &:= |\widehat{A}(x) \cap \widehat{A}(y_j)| \ , \\ H &:= \Big|\widehat{A}(x) \setminus \bigcup_{j\in[L+1]} \widehat{A}(y_j)\Big| \ , \\ Z(x; y_1, \dots, y_{L+1}) &:= \frac{\sum_{j\in[L+1]} \ell_j + \gamma n + H}{(L+1)n} \ . \end{aligned} \tag{10}$$

The numerator counts the existing agreements separately for each slope, together with one additional agreement for each coordinate available for completion. Indeed, the complement of $\bigcup_j(\widehat{A}(x) \cap \widehat{A}(y_j))$ has size $\gamma n + H$. If every $\ell_j < n(1-p)$, the total deficit from the target $n(1-p)$ is

$$\sum_{j=1}^{L+1} \big(n(1-p) - \ell_j\big) \ .$$

This deficit can be filled by assigning disjoint sets of available coordinates to the slopes exactly when $Z \geq 1-p$. Thus, for pairwise distinct slopes with overlaps below the target, list-decodability forces $Z < 1-p$. We establish these conditions and give the completion argument in Section 8.1.

To average the score, sample $\mathcal{X}$ uniformly from B' and, conditionally on $\mathcal{X}$, sample $Y_1, \dots, Y_{L+1}$ independently and uniformly from $\mathsf{B}' \setminus \{\mathcal{X}\}$. Each overlap has expected size $n\hat{\mu}_2$, while $\mathbb{E}[H] = n\nu_{L+1}$ by Definition 5.5. Consequently,

$$\mathcal{Z}_{L+1} := \mathbb{E}[Z(\mathcal{X}; Y_1, \dots, Y_{L+1})] = \hat{\mu}_2 + \frac{\gamma + \nu_{L+1}}{L+1} \ . \tag{11}$$

The general set estimates therefore give the lower bound we will compare with list-decodability.

Corollary 5.8 (Lower bound on the expected completion score)**.** *With the notation above,*

$$\mathcal{Z}_{L+1} \ \geq \ F_{L+1}(\gamma) - \frac{(1-\gamma)\gamma}{M-1} \ .$$

Proof. Apply the third inequality of Lemma 5.6 to the trimmed family with $k = L+1$, and use (11). □

6 List-decoding families of linear maps over prime fields

Fix a prime ℓ, and let V, W be finite-dimensional $\mathbb{F}_\ell$-vector spaces. We regard a linear map $A : V \to W$ as a word indexed by V with alphabet W. For linear A, B, rank-nullity gives

$$\mathsf{agr}(A, B) = \Pr_{z \in V}[(A - B)(z) = 0] = \ell^{-\mathsf{rank}(A-B)} \ . \tag{12}$$

We show that minimum rank distance three forces bounded Hamming lists at every agreement strictly above ℓ^{-2}. The list bound depends on ℓ and the agreement margin, but not on the dimensions of V and W.

6.1 A sunflower bound with repetitions

The distinct-set bound used below is the classical sunflower lemma [ER60]; the version allowing repetitions follows by bounding their multiplicities.

Lemma 6.1 (Sunflowers in sequences)**.** *Let $k \geq 1$ and $\nu \geq 2$ be integers and set*

$$\mathsf{S}_k(\nu) := k!\,(\nu - 1)^{k+1} + 1 \ .$$

Every sequence of at least $\mathsf{S}_k(\nu)$ sets of size k contains ν terms whose pairwise intersections all coincide. Repeated identical terms are allowed.

Proof. We first recall that a family of *distinct* k-element sets containing no ν-sunflower has size at most $k!(\nu - 1)^k$, by induction on k, the case $k = 0$ being the bound 1. Given such a family, a maximal pairwise disjoint subfamily has at most $\nu - 1$ members, so its union Y has size at most $k(\nu - 1)$ and meets every member of the family. Some element of Y therefore lies in at least a $1/[k(\nu - 1)]$ fraction of the family; deleting it from the sets containing it produces distinct $(k - 1)$-element sets with no ν-sunflower, and the induction hypothesis gives the claim.

Now let a sequence of at least $k!(\nu - 1)^{k+1} + 1$ sets of size k be given. If some set occurs at least ν times in the sequence, those ν terms have all pairwise intersections equal to that set. Otherwise every multiplicity is at most $\nu - 1$, so the number of distinct sets occurring is more than $k!(\nu - 1)^k$, and the family of distinct sets contains a ν-sunflower; its petals have all pairwise intersections equal to its core. □

For a three-dimensional space over $\mathbb{F}_\ell$, the set of nonzero vectors has size $k_\ell := \ell^3 - 1$. The relevant sunflower threshold is therefore

$$\mathsf{S}_{k_\ell}(\nu) = k_\ell!\,(\nu - 1)^{\ell^3} + 1 \ . \tag{13}$$

6.2 Bounding the close rank-three maps

For $0 < \varepsilon \leq 1 - \ell^{-2}$, define

$$\begin{aligned} h_{\ell,\varepsilon} &:= \left\lceil \frac{4\ell^6}{\varepsilon^2} \right\rceil \ , \\ T_\ell &:= 2^{\ell^6} \ , \\ D_{\ell,\varepsilon} &:= \mathsf{S}_{k_\ell}(\mathsf{S}_{k_\ell}(T_\ell h_{\ell,\varepsilon})) - 1 = O_\ell\Big(\varepsilon^{-2\ell^6}\Big) \ . \end{aligned} \tag{14}$$

Lemma 6.2 (Close rank-three maps)**.** *Let $A_1, \ldots, A_N : V \to W$ be $\mathbb{F}_\ell$-linear maps of rank three, with $\mathsf{rank}(A_i - A_j) \geq 3$ for $i \neq j$. For every function $f : V \to W$, at most $D_{\ell,\varepsilon}$ of these maps satisfy $\mathsf{agr}(A_i, f) \geq \ell^{-2} + \varepsilon$.*

Proof. Suppose at least $D_{\ell,\varepsilon}+1$ maps satisfy the agreement condition, and discard the others. Write

$$U_i := \mathrm{Im}(A_i^\vee) \subseteq V^\vee \ , \qquad W_i := \mathrm{Im}(A_i) \subseteq W \ .$$

Both spaces have dimension three, so $U_i \setminus \{0\}$ and $W_i \setminus \{0\}$ have k_ℓ elements. Apply Lemma 6.1 first to the row spaces, retaining $\mathsf{S}_{k_\ell}(T_\ell h_{\ell,\varepsilon})$ maps, and then to their image spaces, retaining $T_\ell h_{\ell,\varepsilon}$ maps. There are subspaces $U \subseteq V^\vee$ and $W_0 \subseteq W$ such that

$$U_i \cap U_j = U \ , \qquad W_i \cap W_j = W_0 \qquad (i \neq j) \ . \tag{15}$$

Put $r := \dim U$ and $w := \dim W_0$, so $0 \leq r, w \leq 3$.

For $z \in V$, let $\pi(z) \in U^\vee$ be the functional $\pi(z)(\lambda) := \lambda(z)$. The map $\pi : V \to U^\vee$ is linear and onto, so uniform $z \in V$ gives uniform $\tau := \pi(z) \in U^\vee$. Conditional on τ, each $A_i(z)$ is uniform on an affine subspace of W_i of size ℓ^{3-r}. Indeed, fixing τ fixes the values of the r common row functionals, while the remaining $3-r$ row coordinates remain uniform. Its conditional point probabilities consequently lie in $\{0, \beta\}$, where $\beta := \ell^{r-3}$.

For $i \neq j$, the values $A_i(z)$ and $A_j(z)$ are conditionally independent given τ. To see this, $\dim(U_i + U_j) = 6 - r$ by (15). After fixing the r common coordinates, the remaining $3-r$ coordinates from each row space are jointly uniform and independent.

Consider the tables

$$p_i(\tau, y) := \Pr_{z \in V}[A_i(z) = y \mid \pi(z) = \tau] \qquad ((\tau, y) \in U^\vee \times W_0) \ .$$

Each has $\ell^{r+w} \leq \ell^6$ entries, all in $\{0, \beta\}$, so there are at most T_ℓ possible tables. Retain $h := h_{\ell,\varepsilon}$ maps sharing a table p, and put

$$\mu := \mathbb{E}_{\tau \in U^\vee}\left[\max_{y \in W_0} p(\tau, y)\right] \leq \beta \ .$$

A common output of two selected maps must lie in W_0. Conditional independence and $p(\tau, y)^2 = \beta p(\tau, y)$ therefore give, for $i \neq j$,

$$\begin{aligned} \ell^{-3} &\geq \Pr_{z \in V}[A_i(z) = A_j(z)] \\ &= \mathbb{E}_{\tau \in U^\vee}\left[\sum_{y \in W_0} p(\tau, y)^2\right] \\ &= \beta\, \mathbb{E}_{\tau \in U^\vee}\left[\sum_{y \in W_0} p(\tau, y)\right] \geq \beta\mu \ . \end{aligned}$$

Together with $\mu \leq \beta$, this yields

$$\mu \leq \min\{\ell^{r-3}, \ell^{-r}\} \leq \ell^{-2} \ , \tag{16}$$

where the last inequality holds for each $r \in \{0, 1, 2, 3\}$.

For $y \in W$, let $N_y(z) := |\{i \in [h] : A_i(z) = y\}|$. For $y \in W_0$, conditional pairwise independence gives conditional mean $hp(\tau, y)$ and variance at most $h/4$. Thus

$$\mathbb{E}_{z \in V}\left[\max_{y \in W_0} N_y(z) \mid \pi(z) = \tau\right] \leq h \max_{y \in W_0} p(\tau, y) + |W_0| \frac{\sqrt{h}}{2}$$

$$\leq h \max_{y \in W_0} p(\tau, y) + \frac{\ell^3 \sqrt{h}}{2} \ .$$

For $y \notin W_0$, the pairwise image intersections force $N_y(z) \leq 1$. Averaging over τ, using (16), and then using the choice of h gives

$$\begin{aligned} \frac{1}{h} \sum_{i=1}^{h} \mathsf{agr}(A_i, f) &\leq \frac{1}{h} \, \mathbb{E}_{z \in V} \left[\max_{y \in W} N_y(z) \right] \\ &\leq \mu + \frac{\ell^3}{2\sqrt{h}} + \frac{1}{h} \\ &\leq \ell^{-2} + \frac{\varepsilon}{4} + \frac{\varepsilon^2}{4\ell^6} < \ell^{-2} + \varepsilon \ . \end{aligned}$$

This contradicts the agreement condition on the selected maps. □

6.3 The list bound

The graph reduction below is related to the deletion lemma in [GGR09]. That work also uses low-rank reductions for list-decoding linear maps; the rank-three estimate needed here is supplied by Lemma 6.2.

Proposition 6.3 (List-decoding at agreement above ℓ^{-2})**.** *Let $\mathcal{H} \subseteq \mathsf{Hom}_{\mathbb{F}_\ell}(V, W)$ satisfy $\mathsf{rank}(A - B) \geq 3$ for distinct $A, B \in \mathcal{H}$. For every $0 < \varepsilon \leq 1 - \ell^{-2}$, the code $\mathcal{H}$ is $(1 - \ell^{-2} - \varepsilon, \Lambda_{\ell,\varepsilon})$-list-decodable, where*

$$\Lambda_{\ell,\varepsilon} := \left\lceil \frac{(\ell - 1) D_{\ell,\varepsilon} + \ell^4 - 1}{\ell^4 (\ell^{-2} + \varepsilon)^2 - 1} \right\rceil = O_\ell\Big(\varepsilon^{-(2\ell^6 + 1)} \Big) \ . \tag{17}$$

Proof. Fix $f : V \to W$ and distinct $A_1, \ldots, A_M \in \mathcal{H}$ with agreement at least $\alpha := \ell^{-2} + \varepsilon$ with f. Join i, j by an edge when $\mathsf{rank}(A_i - A_j) = 3$. Every vertex has degree at most $D_{\ell,\varepsilon}$: subtracting A_i from each of its neighbors produces rank-three maps at pairwise rank distance at least three, each agreeing with $f - A_i$ on at least an α fraction of the coordinates, so Lemma 6.2 applies.

Maps joined by an edge have agreement ℓ^{-3}, and maps not joined by an edge have agreement at most ℓ^{-4}. There are at most $M D_{\ell,\varepsilon}$ ordered edges. Hence

$$\sum_{i \neq j} \Pr_{z \in V}[A_i(z) = A_j(z)] \leq \frac{M(M-1) + (\ell - 1) M D_{\ell,\varepsilon}}{\ell^4} \ .$$

Let $b(z) := |\{i \in [M] : A_i(z) = f(z)\}|$. Then $\mathbb{E}_{z \in V}[b(z)] \geq M\alpha$ and $b(z) \leq M$. Since $\mathbb{E}[b(z)^2] \geq (\mathbb{E}[b(z)])^2$, and $b(z)(b(z) - 1)$ counts ordered pairs of maps both equal to $f(z)$, we obtain

$$\begin{aligned} M^2\alpha^2 - M &\leq \mathbb{E}_{z \in V}[b(z)(b(z) - 1)] \\ &= \sum_{i \neq j} \Pr_{z \in V}[A_i(z) = A_j(z) = f(z)] \\ &\leq \frac{M(M-1) + (\ell - 1) M D_{\ell,\varepsilon}}{\ell^4} \ . \end{aligned}$$

For $M > 0$, this rearranges to

$$M(\ell^4 \alpha^2 - 1) \leq (\ell - 1) D_{\ell,\varepsilon} + \ell^4 - 1 \ .$$

Finally, $\ell^4(\ell^{-2} + \varepsilon)^2 - 1 = 2\ell^2\varepsilon + \ell^4\varepsilon^2 \geq 2\ell^2\varepsilon$. Combining this with (14) gives the stated polynomial bound. □

7 A unified MCA bound for polynomial generators

Throughout this section, let Σ be a finite-dimensional $\mathbb{F}$-vector space. We fix an $\mathbb{F}$-linear code $\mathcal{C} \subseteq \Sigma^n$ of minimum distance $\Delta_{\mathcal{C}}$ and relative distance $\delta_{\mathcal{C}} := \Delta_{\mathcal{C}}/n \in (0,1]$, a degree $d \geq 1$, and a nonempty seed set $S \subseteq \mathbb{F}$. We write

$$G_d \colon S \to \mathbb{F}^{d+1} \ ,$$
$$G_d(x) := (1, x, x^2, \ldots, x^d)$$

for the univariate powers generator, whose output size is $\ell := d+1$. Put $h := d+2 = \ell + 1$, so that its generalized Johnson radius is $J_h(\delta_{\mathcal{C}})$, as defined in Section 3.1. We use the radius γ and the integer-radius convention from Section 4 throughout this section.

Theorem 7.1 (Explicit gap-dependent bound)**.** *Let Σ be a finite-dimensional $\mathbb{F}$-vector space and let $\mathcal{C} \subseteq \Sigma^n$ be an $\mathbb{F}$-linear code with relative distance $\delta_{\mathcal{C}}$. Let $d \geq 1$ and $\gamma \in [0,1]$. With the notation of* (2)*, if*

$$2\gamma n < \Delta_{\mathcal{C}}$$
$$\gamma < J_h(\delta_{\mathcal{C}}) \ ,$$

then G_d has mutual correlated agreement for $\mathcal{C}$ with error

$$\epsilon_{\mathsf{MCA}}(\gamma) \ \leq \ \xi_{\mathcal{C},d,|S|}(\gamma) := \frac{15\,d}{|S|} \cdot \max\left\{1,\ \gamma n,\ \frac{\gamma}{J_h(\delta_{\mathcal{C}}) - \gamma}\right\} \ .$$

The constant is independent of d, n, $|S|$, $\delta_{\mathcal{C}}$, γ and $\dim_{\mathbb{F}} \Sigma$. For arbitrary real γ, replace γ throughout the hypotheses and the bound by $\lfloor \gamma n \rfloor / n$, as in Section 4.

Proof. Fix words $\mathbf{u}_0, \ldots, \mathbf{u}_d \in \Sigma^n$ and let B be as in Definition 4.1. It suffices to prove

$$|\mathsf{B}| \ \leq \ 15\,d \cdot m \ ,$$
$$m := \max\left\{1, \gamma n, \tfrac{\gamma}{J_h(\delta_{\mathcal{C}}) - \gamma}\right\} \ , \tag{18}$$

and divide by $|S|$. If $\gamma = 0$ this follows from Lemma 4.3, so assume $\gamma n \geq 1$. Set

$$\tau \ := \ \min\left\{\frac{1 - \gamma + \sqrt{1 - \delta_{\mathcal{C}}}}{2},\ 1 - \delta_{\mathcal{C}} + \gamma\right\} \ ,$$

and note that $(1 - \delta_{\mathcal{C}}) < \tau \leq \frac{1-\gamma+\sqrt{1-\delta_{\mathcal{C}}}}{2} < (1-\gamma)$ and $1 - \gamma - \tau \geq \frac{1-\gamma-\sqrt{1-\delta_{\mathcal{C}}}}{2} > 0$, by (3). Recall that we say that a curve P is *large* if $z_P > \tau$ (see Definition 4.4).

If $|\mathsf{B}| \leq d(\gamma n + 1)$ then $|\mathsf{B}| \leq 2d\gamma n \leq 2dm$ which finishes the proof. So assume $|\mathsf{B}| > d(\gamma n + 1)$. By Lemma 4.5, no curve can be such that all of B is on it; hence by Lemma 4.10 every curve such that at least $d+1$ seeds of B lie on it satisfies $z_P \leq 1 - \delta_{\mathcal{C}} + \gamma$.

If $\tau = 1 - \delta_{\mathcal{C}} + \gamma$, then by the preceding paragraph no curve interpolating ℓ distinct seeds of B is large; we set $\mathsf{B}' := \mathsf{B}$ and $L := 0$. Otherwise $\tau = \frac{1-\gamma+\sqrt{1-\delta_{\mathcal{C}}}}{2} < 1 - \delta_{\mathcal{C}} + \gamma$; by Lemma 4.11 there are at most $L \leq 4$ large curves, and we let $\mathsf{B}' \subseteq \mathsf{B}$ be the set of seeds that do not lie on any large curve, so that, by Lemma 4.5,

$$|\mathsf{B}| \ \leq \ L\,d(\gamma n + 1) + M \ ,$$

$$M := |\mathsf{B}'| \ .$$

In both cases, every curve interpolating ℓ distinct seeds of B' satisfies $z_P \le \tau$.

We claim that in the second case $\gamma n \ge 2$. Indeed, $\frac{1-\gamma+\sqrt{1-\delta_{\mathcal{C}}}}{2} < 1-\delta_{\mathcal{C}}+\gamma$ reads $3\gamma > 1+\sqrt{1-\delta_{\mathcal{C}}}-2(1-\delta_{\mathcal{C}}) = (1+2\sqrt{1-\delta_{\mathcal{C}}})(1-\sqrt{1-\delta_{\mathcal{C}}})$, while $\Delta_{\mathcal{C}} \ge 2\gamma n+1$ gives $\delta_{\mathcal{C}} \ge \gamma\,\frac{2\gamma n+1}{\gamma n}$, i.e. $\gamma \le \frac{\gamma n(1-\sqrt{1-\delta_{\mathcal{C}}})(1+\sqrt{1-\delta_{\mathcal{C}}})}{2\gamma n+1}$. Combining, $(1+2\sqrt{1-\delta_{\mathcal{C}}})(2\gamma n+1) < 3\gamma n(1+\sqrt{1-\delta_{\mathcal{C}}})$, i.e. $\gamma n(1-\sqrt{1-\delta_{\mathcal{C}}}) > 1+2\sqrt{1-\delta_{\mathcal{C}}}$, so $\gamma n > \frac{1+2\sqrt{1-\delta_{\mathcal{C}}}}{1-\sqrt{1-\delta_{\mathcal{C}}}} \ge 1$.

If $M < 6\ell$ then $M \le 6\ell - 1 = 6d+5$. Assume $M \ge 6\ell$.

Let P interpolate ℓ distinct seeds of B'. Then $z_P \le \tau < (1-\gamma)$, so by Lemma 4.5 at most $d + \frac{d\gamma}{1-\gamma-\tau}$ seeds of B' lie on P. Put

$$K \ := \ \frac{d\gamma}{1-\gamma-\tau} \ .$$

The $\ell = d+1$ interpolation seeds are already on P, leaving at most $K-1$ further seeds. We use the upper bound K below.

Sample $h = \ell+1$ distinct seeds $x_1, \dots, x_h$ of B' uniformly without replacement and condition on $x_1, \dots, x_\ell$; let P be their interpolating curve, whose domain T_P is their common agreement set (Lemma 4.6). If x_h lies on P then $T_P \subseteq A(x_h)$ — for $i \in T_P$ we have $\mathbf{u}_i(x_h) = P_i(x_h) = \mathbf{c}_{x_h}[i]$ — so the common agreement fraction of all h seeds is exactly z_P; otherwise the tuple is noncoherent and that fraction is at most $1-\delta_{\mathcal{C}}$ by Lemma 4.9. By the bound on the number of further seeds on P, x_h lies on P with conditional probability at most $\frac{K}{M-\ell}$, so the conditional expectation of the common agreement fraction is at most

$$(1-\delta_{\mathcal{C}}) + \frac{K}{M-\ell}\,(z_P - (1-\delta_{\mathcal{C}}))^+ \ .$$

The function $y \mapsto (y-(1-\delta_{\mathcal{C}}))^+$ is convex and vanishes at $y=0$, hence lies below the chord through $(0,0)$ and $(\tau, \tau-(1-\delta_{\mathcal{C}}))$ on $[0,\tau]$; since $0 \le z_P \le \tau$ this gives $(z_P-(1-\delta_{\mathcal{C}}))^+ \le \frac{\tau-(1-\delta_{\mathcal{C}})}{\tau} z_P$. Averaging over $x_1, \dots, x_\ell$ and using $\mathbb{E}[z_P] = \mu_\ell$,

$$\begin{aligned} \mu_h \ &\le \ (1-\delta_{\mathcal{C}}) + \frac{d\gamma\, X}{M-\ell}\,\mu_\ell \ , \\ X &:= \frac{\tau-(1-\delta_{\mathcal{C}})}{\tau\,(1-\gamma-\tau)} \ . \end{aligned} \tag{19}$$

Let $v := \mu_\ell^{1/\ell} \in [0,1]$. Since $M \ge 6\ell$ and $(1-\gamma) > 1/2$, Corollary 5.3 with $j = \ell$ gives $v \ge 1-\gamma-\theta > 0$ with $\theta = \frac{d\gamma}{M-d}$. Combining (19) with Lemma 5.4, multiplying by $M-\ell$ and dividing by $v^\ell > 0$,

$$\begin{aligned} Mv - \ell \le \ (M-\ell)\frac{(1-\delta_{\mathcal{C}})}{v^\ell} + d\gamma X \ , \\ (M-\ell)\left(v - \frac{(1-\delta_{\mathcal{C}})}{v^\ell}\right) \le \ \ell(1-v) + d\gamma X \ . \end{aligned}$$

Since $v \le 1$, using $\sum_{m=0}^{h-1}(1-\delta_{\mathcal{C}})^{m/h} = \delta_{\mathcal{C}}/J_h(\delta_{\mathcal{C}})$,

$$v - \frac{1-\delta_{\mathcal{C}}}{v^\ell} = \frac{v^h - (1-\delta_{\mathcal{C}})}{v^{h-1}}$$

$$= \left(v - (1-\delta_{\mathcal{C}})^{1/h}\right) \sum_{m=0}^{h-1} \left(\frac{(1-\delta_{\mathcal{C}})^{1/h}}{v}\right)^m$$
$$\geq \frac{\delta_{\mathcal{C}}}{J_h(\delta_{\mathcal{C}})} \left(v - (1-\delta_{\mathcal{C}})^{1/h}\right) \text{ whenever } v \geq (1-\delta_{\mathcal{C}})^{1/h} \ .$$

If instead $v < (1-\delta_{\mathcal{C}})^{1/h}$, then $\frac{\delta_{\mathcal{C}}}{J_h(\delta_{\mathcal{C}})}(v - (1-\delta_{\mathcal{C}})^{1/h}) < 0 \leq \ell(1-v) + d\gamma X$. In either case

$$(M-\ell)\,\frac{\delta_{\mathcal{C}}}{J_h(\delta_{\mathcal{C}})}\,(v - (1-\delta_{\mathcal{C}})^{1/h}) \ \leq \ \ell(1-v) + d\gamma X \ . \tag{20}$$

Insert $v \geq 1-\gamma-\theta$ and $1-v \leq \gamma+\theta$ into (20):

$$(M-\ell)\,\frac{\delta_{\mathcal{C}}}{J_h(\delta_{\mathcal{C}})}\,(J_h(\delta_{\mathcal{C}}) - \gamma) \ \leq \ \ell\gamma + d\gamma X + \theta\left[\ell + \frac{\delta_{\mathcal{C}}}{J_h(\delta_{\mathcal{C}})}(M-\ell)\right] \ .$$

Because $d < \ell$ we have $\theta \leq \frac{d\gamma}{M-\ell}$, so $\theta\,\frac{\delta_{\mathcal{C}}}{J_h(\delta_{\mathcal{C}})}(M-\ell) \leq \frac{\delta_{\mathcal{C}}}{J_h(\delta_{\mathcal{C}})}\,d\gamma$, and because $M \geq 6\ell$ we have $\frac{\ell}{M-\ell} \leq \frac{1}{5}$, so $\theta\ell \leq \frac{d\gamma}{5}$. Hence

$$(M-\ell)\,\frac{\delta_{\mathcal{C}}}{J_h(\delta_{\mathcal{C}})}\,(J_h(\delta_{\mathcal{C}}) - \gamma) \ \leq \ \gamma\left[\ell + d\left(X + \frac{\delta_{\mathcal{C}}}{J_h(\delta_{\mathcal{C}})} + \tfrac{1}{5}\right)\right] \ .$$

Using $M - \ell \geq \frac{5}{6}M$, $1-\gamma-(1-\delta_{\mathcal{C}})^{1/h} = J_h(\delta_{\mathcal{C}}) - \gamma$ and $\ell/d \leq 2$,

$$\begin{aligned} M &\leq \frac{6}{5}\frac{d\gamma}{J_h(\delta_{\mathcal{C}})-\gamma}\left(1 + \frac{J_h(\delta_{\mathcal{C}})}{\delta_{\mathcal{C}}}\left(\frac{\ell}{d} + \frac{1}{5} + X\right)\right) \\ &\leq \frac{6}{5}\frac{d\gamma}{J_h(\delta_{\mathcal{C}})-\gamma}\left(1 + \frac{J_h(\delta_{\mathcal{C}})}{\delta_{\mathcal{C}}}\left(\frac{11}{5} + X\right)\right) \\ &\leq \frac{6}{5}\cdot\frac{36}{5}\cdot\frac{d\gamma}{J_h(\delta_{\mathcal{C}})-\gamma} \\ &= \frac{216}{25}\frac{d\gamma}{J_h(\delta_{\mathcal{C}})-\gamma} \ . \end{aligned}$$

In the last inequality we used Lemma B.1.

Combining the bounds for $M < 6\ell$ and $M \geq 6\ell$,

$$\begin{aligned} M \ &\leq \ \max\left\{6d+5, \ \tfrac{216}{25}\,d\,\tfrac{\gamma}{J_h(\delta_{\mathcal{C}})-\gamma}\right\} \ , \\ |\mathsf{B}| \ &\leq \ L\,d(\gamma n + 1) + M \ . \end{aligned}$$

Recall $m = \max\{1, \gamma n, \gamma/(J_h(\delta_{\mathcal{C}}) - \gamma)\} \geq \max\{1, \gamma n\}$, so $\frac{216}{25} d \frac{\gamma}{J_h(\delta_{\mathcal{C}})-\gamma} \leq \frac{216}{25}\,dm$ and $6d+5 \leq 11d \leq 11dm$.

- If $L = 0$, then $|\mathsf{B}| = M \leq \max\{11, \ \frac{216}{25}\}\,dm = 11\,dm$.
- If $L \geq 1$, then $\gamma n \geq 2$ as shown when removing large curves, so $Ld(\gamma n + 1) \leq 4d \cdot \frac{3}{2}\gamma n \leq 6dm$ and $6d + 5 \leq \frac{11}{2} d\gamma n \leq \frac{11}{2} dm$; hence $|\mathsf{B}| \leq 6dm + \max\{\frac{11}{2}, \ \frac{216}{25}\}dm \leq \frac{366}{25}\,dm$.

In all cases (18) holds, which completes the proof. □

The factor $\gamma/(J_h(\delta_{\mathcal{C}}) - \gamma)$ is needed only when $n\,(J_h(\delta_{\mathcal{C}}) - \gamma) < 1$ below the generalized Johnson radius. As soon as the error radius stays one coordinate away from that radius, Theorem 7.1 gives the scaling $n\gamma d/|S|$, uniformly in the degree.

Corollary 7.2 (Uniform scaling)**.** *In the setting of Theorem 7.1:*

1. If $n\,(J_h(\delta_{\mathcal{C}}) - \gamma) \geq 1$ — which holds in particular when $\gamma \leq J_h(\delta_{\mathcal{C}}) - 1/n$ — then

$$\epsilon_{\mathsf{MCA}}(\gamma) \leq \frac{15\,d}{|S|} \cdot \max\{1, n\gamma\} \ .$$

2. If $J_h(\delta_{\mathcal{C}}) \geq \delta_{\mathcal{C}}/2$, then for every $\gamma < \delta_{\mathcal{C}}/2$,

$$\epsilon_{\mathsf{MCA}}(\gamma) \leq \frac{24\,d}{|S|} \cdot \max\{1, n\gamma\} \ .$$

In other words: the $O(n\gamma d/|S|)$ scaling holds up to the generalized Johnson radius with a one-coordinate margin whenever that radius is the smaller of the two, and holds in the entire unique decoding radius whenever unique decoding is the smaller of the two. The two cases meet at the crossover $J_h(\delta_{\mathcal{C}}) = \delta_{\mathcal{C}}/2$, which for $d = 1$ (i.e. $h = 3$) is the solution

$$\delta_{\mathcal{C}} = 3 - \sqrt{5}$$

of $1 - (1 - \delta)^{1/3} = \delta/2$, and which increases with d.

Proof. For the first part, $\gamma \leq J_h(\delta_{\mathcal{C}}) - 1/n$ gives $\gamma n \leq n J_h(\delta_{\mathcal{C}}) - 1$, i.e. $n\,(J_h(\delta_{\mathcal{C}}) - \gamma) \geq 1$; then $\gamma/(J_h(\delta_{\mathcal{C}}) - \gamma) \leq \gamma n$ and $m = \max\{1, \gamma n\}$, so Theorem 7.1 applies verbatim.

For the second part, $2\gamma n < \Delta_{\mathcal{C}}$ and the integrality of $\Delta_{\mathcal{C}}$ and γn give $\gamma n \leq \frac{\Delta_{\mathcal{C}} - 1}{2}$, hence

$$n\,(J_h(\delta_{\mathcal{C}}) - \gamma) = n J_h(\delta_{\mathcal{C}}) - \gamma n \ \geq \ \tfrac{\delta_{\mathcal{C}} n}{2} - \gamma n \ = \ \tfrac{\Delta_{\mathcal{C}}}{2} - \gamma n \ \geq \ \tfrac{1}{2} > 0 \ ,$$

so the hypotheses of Theorem 7.1 hold. Repeating its final estimate with $m' := \max\{1, \gamma n\}$ and $\gamma/(J_h(\delta_{\mathcal{C}}) - \gamma) \leq 2\gamma n \leq 2m'$ gives $\frac{216}{25} d \frac{\gamma}{J_h(\delta_{\mathcal{C}}) - \gamma} \leq \frac{432}{25}\, dm'$, hence $|\mathsf{B}| \leq \max\{11, \frac{432}{25}\} dm' = \frac{432}{25}\, dm'$ when $L = 0$ and $|\mathsf{B}| \leq 6dm' + \frac{432}{25}\, dm' = \frac{582}{25}\, dm'$ when $L \geq 1$. □

Remark 7.3. Let $G = (P_1, \ldots, P_\ell)$ on a rectangular domain $\prod_{i=1}^{s} S_i \subseteq \mathbb{F}^s$, where the $P_j \in \mathbb{F}[\mathtt{X}_1, \ldots, \mathtt{X}_s]$ are linearly independent over $\mathbb{F}$. Put $d_i := \max_j \deg_{\mathtt{X}_i} P_j$ and assume $|S_i| \geq d_i + 1$. Applying Corollary 7.2 to each positive degree d_i and using the propagation results of [BCGM26] gives

$$\epsilon_{\mathsf{MCA}}^{\mathcal{C},G}(\gamma) \leq 24 \max\{1, n\gamma\} \sum_{i=1}^{s} \frac{d_i}{|S_i|} \ ,$$

provided $0 \leq \gamma < \delta_{\mathcal{C}}/2$ and $\gamma \leq J_{d_i+2}(\delta_{\mathcal{C}}) - 1/n$ whenever $d_i \geq 1$ and $J_{d_i+2}(\delta_{\mathcal{C}}) < \delta_{\mathcal{C}}/2$, with the same integer-radius convention.

8 A list-size-dependent bound

Theorem 8.1 (List-size-dependent bound)**.** *Let $\mathcal{C} \subseteq \mathbb{F}^n$ be linear and (p, L)-list-decodable, where $L \geq 1$ is an integer, $|\mathbb{F}| > L$ and $pn \in \{0, \ldots, n-1\}$. Let γn be an integer with $0 \leq \gamma \leq p$, and define*

$$\begin{aligned} F_{L+1}(\gamma) &:= (1-\gamma)^2 + \frac{\gamma + (1-\gamma)\gamma^{L+1}}{L+1} \ , \\ G &:= F_{L+1}(\gamma) - (1-p) \ . \end{aligned} \tag{21}$$

If $G > 0$, then

$$\epsilon^{\mathcal{C}}_{\mathsf{MCA}}(\gamma) \ \leq \ \frac{1}{|\mathbb{F}|} \left[\frac{L(|\mathbb{F}|-1)}{|\mathbb{F}|-L} \left(\gamma n + 1\right) + 1 + \max \left\{ \frac{L \cdot \gamma n}{n(p-\gamma)+1}, \ \frac{(1-\gamma)\gamma + \frac{(1-\gamma)\cdot L \cdot \gamma n}{2(n(p-\gamma)+1)}}{G} \right\} \right] \ . \tag{22}$$

In particular, if $|\mathbb{F}| \geq 2L$, then

$$\epsilon^{\mathcal{C}}_{\mathsf{MCA}}(\gamma) \ \leq \ \left(2L + 1 + \frac{L+1}{G}\right) \frac{n\gamma + 1}{|\mathbb{F}|} \ \leq \ \frac{5L(n\gamma+1)}{G\,|\mathbb{F}|} \ . \tag{23}$$

For arbitrary real radii $0 \leq \gamma \leq p < 1$, apply Theorem 8.1 with p and γ rounded down to $\lfloor pn \rfloor / n$ and $\lfloor \gamma n \rfloor / n$, respectively, in both the hypotheses and the bound. The bound in (23) is linear in L for a fixed positive margin G, whereas the bound in [GCXK26] has quadratic dependence on L. For fixed radii with $p < \delta_{\mathcal{C}}$ and $(1-\gamma)^2 > 1-p$, we have $G \geq (1-\gamma)^2 - (1-p) > 0$ uniformly in L, so the linear dependence holds with a constant independent of L. Moreover, the condition $G > 0$ improves on the square-root condition $(1-\gamma)^2 > 1-p$ in [GCXK26] by retaining the positive finite-list correction in F_{L+1}.

Throughout this section we work in the setting of Theorem 8.1. Fix $\mathbf{v}, \mathbf{w} \in \mathbb{F}^n$, write

$$\mathbf{u}(\mathtt{X}) := \mathbf{v} + \mathtt{X}\mathbf{w} \ ,$$

and put $\mathsf{B} := \mathsf{B}_{\mathcal{C}}(\mathbf{v}, \mathbf{w}; \gamma)$. For each $x \in \mathsf{B}$ fix once and for all a witness set T_x and a codeword $\mathbf{c}_x \in \mathcal{C}$ with $\mathbf{c}_x|_{T_x} = \mathbf{u}(x)|_{T_x}$, and let

$$A(x) := \{i \in [n] : \mathbf{c}_x[i] = \mathbf{u}_i(x)\}$$

be the corresponding full agreement set. Then $A(x) \supseteq T_x$, so $|A(x)| \geq n(1-\gamma)$, and $A(x)$ is again a witness for the badness of x: any codeword equal to $\mathbf{v}$ or $\mathbf{w}$ on $A(x)$ would also equal that word on T_x. The witness codeword need not be unique, and γ may exceed the unique decoding radius.

Delete from B every seed lying on a large codeword line. By Lemmas 4.12 and 4.13 at most $A_L(\gamma n + 1)$ seeds are deleted. Write B' for the set of remaining seeds and $M := |\mathsf{B}'|$, so that

$$|\mathsf{B}| \ \leq \ A_L(\gamma n + 1) + M \ . \tag{24}$$

If $M \leq 1$ then (22) follows at once, since the last term there is nonnegative; assume from now on that $M > 1$. Finally, for each $x \in \mathsf{B}'$ fix a subset

$$\begin{aligned} \widehat{A}(x) &\subseteq A(x) \ , \\ |\widehat{A}(x)| &= (1-\gamma)n \ . \end{aligned}$$

These trimmed sets are used only for counting; Lemma 4.12 continues to refer to the full witnesses $A(x)$.

We bound M by comparing two estimates for the expected completion score from Section 5.3. For a fixed seed, the codeword witnesses at the other seeds determine slopes in $\mathcal{C}$. Completing their agreements with $\mathbf{w}$ shows that $L+1$ distinct slopes must have completion score below $1-p$, since otherwise they would lie in a single decoding ball of radius p. The moment lower bound uses independently sampled neighbors, whose slopes may repeat. We therefore couple this sample to one with distinct slopes. Root counting bounds the size of each slope class, and hence the expected change in the score. Comparing the resulting upper bound with the moment lower bound gives the desired estimate for M.

8.1 List-decodability bounds the completion score

Lemma 8.2 (Slopes)**.** *Let $x \in \mathsf{B}'$ and let $y \in \mathsf{B}' \setminus \{x\}$. Put*

$$d_{x,y} := \frac{\mathbf{c}_y - \mathbf{c}_x}{y - x} \in \mathcal{C} \ ,$$
$$Q_{x,y}(\mathtt{X}) := \mathbf{c}_x + (\mathtt{X} - x) d_{x,y} \ , \tag{25}$$

so that $Q_{x,y}$ is a codeword line through $(x, \mathbf{c}_x)$ and $(y, \mathbf{c}_y)$; we call $d_{x,y}$ the slope *of y relative to x. Then*

1. *$\widehat{A}(x) \cap \widehat{A}(y) \subseteq T_{Q_{x,y}}$, and $d_{x,y}$ agrees with $\mathbf{w}$ on $\widehat{A}(x) \cap \widehat{A}(y)$;*
2. *$Q_{x,y}$ is not large, and consequently*

$$|\widehat{A}(x) \cap \widehat{A}(y)| \ < \ n(1-p) \ ; \tag{26}$$

3. *for every $d \in \mathcal{C}$,*

$$|\{y \in \mathsf{B}' \setminus \{x\} \ : \ d_{x,y} = d\}| \ \leq \ R = \frac{\gamma n}{n(p-\gamma)+1} \ . \tag{27}$$

Proof. (i) For $i \in \widehat{A}(x) \cap \widehat{A}(y)$ the affine polynomial $\mathbf{u}_i - Q_{x,y}[i]$ (where we denote $Q_{x,y}[i] = \mathbf{c}_x[i] + (\mathtt{X} - x) d_{x,y}[i]$) vanishes at the two distinct seeds x and y, hence vanishes identically, i.e. $i \in T_{Q_{x,y}}$. Comparing the coefficients of $\mathtt{X}$ then gives $\mathbf{w}[i] = d_{x,y}[i]$.

(ii) Since B' consists only of seeds that do not lie on any large codeword line, and $Q_{x,y}(x) = \mathbf{c}_x$, the line $Q_{x,y}$ is not large, i.e. $|T_{Q_{x,y}}| < n(1-p)$; now apply (i).

(iii) If no y has slope d there is nothing to prove, so let y have slope d. All such y lie, together with x itself, on the single codeword line $Q := \mathbf{c}_x + (\mathtt{X} - x)d = Q_{x,y}$, which is not large by (ii); thus $s_Q \geq pn + 1 > \gamma n$. By Lemma 4.12,

$$|\{y : d_{x,y} = d\}| \ \leq \ m_Q - 1 \ \leq \ \frac{s_Q}{s_Q - \gamma n} - 1 = \frac{\gamma n}{s_Q - \gamma n} \ \leq \ R \ .$$

□

Lemma 8.3 (Distinct slopes force a low score)**.** *If the slopes $d_{x,y_1}, \ldots, d_{x,y_{L+1}}$ are pairwise distinct, then*

$$Z(x; y_1, \ldots, y_{L+1}) \ < \ 1 - p \ . \tag{28}$$

Proof. Suppose not. By (26) the deficits $e_j := n(1-p) - \ell_j$ are positive integers, and $Z \geq 1-p$ reads $\sum_j \ell_j \geq (L+1)(n(1-p)) - \gamma n - H$, i.e. $\sum_j e_j \leq \gamma n + H$. The set

$$\mathcal{D} := \Big([n] \setminus \widehat{A}(x)\Big) \cup \Big(\widehat{A}(x) \setminus \bigcup_{j \in [L+1]} \widehat{A}(y_j)\Big)$$

has size $\gamma n + H$ and is disjoint from every $\widehat{A}(x) \cap \widehat{A}(y_j)$, so we may choose pairwise disjoint $E_j \subseteq \mathcal{D}$ with $|E_j| = e_j$. Let $z \in \mathbb{F}^n$ equal $\mathbf{w}$ on $\bigcup_j (\widehat{A}(x) \cap \widehat{A}(y_j))$, equal d_{x,y_j} on E_j, and be arbitrary elsewhere. By Lemma 8.2, d_{x,y_j} agrees with z on $\widehat{A}(x) \cap \widehat{A}(y_j)$ and on E_j, hence on at least $\ell_j + e_j = n(1-p)$ coordinates, i.e. $\Delta(d_{x,y_j}, z) \leq pn$. Thus $L+1$ distinct codewords lie within relative distance p of z, contradicting list-decodability. □

8.2 Distinct-slope sampling and the upper moment bound

The independent neighbors $Y_1, \ldots, Y_{L+1}$ need not have distinct slopes, so Lemma 8.3 does not apply to them directly. We couple them to a distinct-slope tuple and bound the change in the score using the following lemma.

Lemma 8.4 (Sensitivity of the score)**.** *For every $x \in \mathsf{B}'$ and all neighbors $y_1, \ldots, y_{L+1}$,*

$$Z(x; y_1, \ldots, y_{L+1}) = \frac{\gamma}{L+1} + \frac{1}{(L+1)n} \sum_{i \in \widehat{A}(x)} \max\{r_i, 1\} \ ,$$

$$r_i := |\{j \in [L+1] : i \in \widehat{A}(y_j)\}| \ . \tag{29}$$

Consequently, changing one of the $L+1$ neighbors changes Z by at most $(1-\gamma)/(L+1)$.

Proof. We have $\sum_j \ell_j = \sum_{i \in \widehat{A}(x)} r_i$ and $H = |\{i \in \widehat{A}(x) : r_i = 0\}|$, so $\sum_j \ell_j + H = \sum_{i \in \widehat{A}(x)} \max\{r_i, 1\}$; together with $\gamma n/((L+1)n) = \gamma/(L+1)$ this is (29). Replacing one neighbor changes each r_i by at most one, hence each summand by at most one, and there are $|\widehat{A}(x)| = (1-\gamma)n$ summands. □

Lemma 8.5 (Upper bound on the expected score)**.** *If $M - 1 > LR$, then*

$$\mathcal{Z}_{L+1} \ < \ 1 - p + \frac{(1-\gamma)LR}{2(M-1)} \ . \tag{30}$$

Proof. Condition on $\mathcal{X} = x$, so that there are $M-1$ neighbors. By Lemma 8.2(iii) each slope class contains at most R neighbors, so any L slope classes together contain at most $LR < M-1$ of them; in particular the neighbors meet at least $L+1$ distinct slope classes.

We couple $Y_1, \ldots, Y_{L+1}$ to neighbors $\widetilde{Y}_1, \ldots, \widetilde{Y}_{L+1}$ with pairwise distinct slopes, one step at a time. At step j, keep $\widetilde{Y}_j := Y_j$ if the slope of Y_j differs from those of $\widetilde{Y}_1, \ldots, \widetilde{Y}_{j-1}$; otherwise let $\widetilde{Y}_j$ be a uniformly sampled neighbor lying outside those $j-1$ slope classes. The latter set is nonempty, because the excluded classes contain at most $(j-1)R \leq LR < M-1$ neighbors. Since Y_j is uniform on the neighbors and independent of $\widetilde{Y}_1, \ldots, \widetilde{Y}_{j-1}$,

$$\Pr\left[\widetilde{Y}_j \neq Y_j \,\middle|\, \mathcal{X}, \widetilde{Y}_1, \ldots, \widetilde{Y}_{j-1}\right] \ \leq \ \frac{(j-1)R}{M-1} \ .$$

Writing $Z := Z(\mathcal{X}; Y_1, \ldots, Y_{L+1})$ and $\widetilde{Z} := Z(\mathcal{X}; \widetilde{Y}_1, \ldots, \widetilde{Y}_{L+1})$ and changing the neighbors one at a time, Lemma 8.4 gives $\left|Z - \widetilde{Z}\right| \leq \frac{1-\gamma}{L+1} |\{j : \widetilde{Y}_j \neq Y_j\}|$, whence

$$\mathbb{E}\left[\left|Z - \widetilde{Z}\right|\right] \ \leq \ \frac{1-\gamma}{L+1} \sum_{j=1}^{L+1} \frac{(j-1)R}{M-1} = \frac{(1-\gamma)LR}{2(M-1)} \ .$$

The tuple $\widetilde{Y}_1, \ldots, \widetilde{Y}_{L+1}$ has distinct slopes, so $\widetilde{Z} < 1 - p$ by Lemma 8.3. We obtain that $\mathcal{Z}_{L+1} = \mathbb{E}[Z] \leq \mathbb{E}\left[\widetilde{Z}\right] + \mathbb{E}\left[\left|Z - \widetilde{Z}\right|\right] < 1 - p + \frac{(1-\gamma)LR}{2(M-1)}$. □

8.3 Proof of Theorem 8.1

Proof. We may assume $M > 1$, as noted after (24). If $M - 1 \le LR$ then $M \le 1 + LR$. Otherwise Lemma 8.5 applies and, combined with Corollary 5.8,

$$F_{L+1}(\gamma) - \frac{(1-\gamma)\gamma}{M-1} \;\le\; \mathcal{Z}_{L+1} \;<\; 1 - p + \frac{(1-\gamma)LR}{2(M-1)} \;\;,$$
$$\text{i.e., } G \;<\; \frac{(1-\gamma)\gamma + (1-\gamma)LR/2}{M-1} \;\;.$$

In either case

$$M \;\le\; 1 + \max\left\{LR,\; \frac{(1-\gamma)\gamma + (1-\gamma)LR/2}{G}\right\} \;\;,$$

and adding the at most $A_L(\gamma n + 1)$ deleted seeds, by (24) and Lemma 4.13, gives (22) after dividing by $|\mathbb{F}|$ and maximizing over $\mathbf{v}, \mathbf{w}$.

For (23), note first that $G \le p \le 1$, since $F_{L+1}(\gamma) \le 1$: indeed, for $L \ge 1$ and $\gamma^{L+1} \le \gamma$,

$$F_{L+1}(\gamma) \;\le\; (1-\gamma)^2 + \frac{\gamma(1 + (1-\gamma))}{2} = (1-\gamma)^2 + \frac{\gamma(2-\gamma)}{2} = 1 - \gamma + \frac{\gamma^2}{2} \;\le\; 1 \;\;.$$

Now $R \le \gamma n$, $(1-\gamma)\gamma \le \gamma \le \gamma n$ (as $n \ge 1$) and $(1-\gamma) \le 1$ give

$$\max\left\{LR,\; \frac{(1-\gamma)\gamma + (1-\gamma)LR/2}{G}\right\} \;\le\; \max\left\{L\gamma n,\; \frac{\gamma n + L\gamma n/2}{G}\right\} \;\le\; \frac{(L+1)\gamma n}{G} \;\;,$$

using $G \le 1$ in the first entry and $1 + L/2 \le L + 1$ in the second. Since $A_L \le 2L$ for $|\mathbb{F}| \ge 2L$ and $1 \le \gamma n + 1$ and $\gamma n \le \gamma n + 1$, the bracket in (22) is at most $\left(2L + 1 + \frac{L+1}{G}\right)(\gamma n + 1)$. Finally $2L + 1 \le 3L$ and $L + 1 \le 2L$ for $L \ge 1$, so, again by $G \le 1$,

$$2L + 1 + \frac{L+1}{G} \;\le\; \frac{3L}{G} + \frac{2L}{G} = \frac{5L}{G} \;\;.$$

□

9 An obstruction to uniform exponents above one half

Theorem 9.1 (An obstruction to uniform exponents above $1/2$)**.** *Fix a prime ℓ and $0 < \eta < (\ell^{-1} - \ell^{-2})/2$. There is an integer $m_0 = O_\ell(\log(2/\eta))$ such that, for every $m \geq m_0$ and $|\mathbb{F}| = \ell^m$, there exist an $\mathbb{F}$-linear code $\mathcal{C} \subseteq \mathbb{F}^n$ and words $\mathbf{u}, \mathbf{v} \in \mathbb{F}^n$ such that*

$$\begin{aligned}
&n = O_\ell\Big(\eta^{-(2\ell^6+4)} m^2\Big) \ , \\
&\mathcal{C} \text{ is } (1 - \ell^{-2} - \eta, L_{\ell,\eta})\text{-list-decodable} \ , \\
&L_{\ell,\eta} = O_\ell\Big(\eta^{-(2\ell^6+2)}\Big) \ , \\
&\delta(\mathbf{u}, \mathcal{C}) \geq 1 - \ell^{-2} - \eta \ , \\
&\delta(\mathbf{u} + x\mathbf{v}, \mathcal{C}) \leq 1 - \ell^{-1} + \eta \qquad (x \in \mathcal{P}) \ .
\end{aligned} \tag{31}$$

where

$$\mathcal{P} := \{a^{\ell-1} : a \in \mathbb{F}^\times\} \ , \qquad |\mathcal{P}| = \frac{|\mathbb{F}| - 1}{\ell - 1} \ . \tag{32}$$

The implicit constants depend only on ℓ, and the sampled evaluation coordinates can be chosen distinct. Consequently,

$$\epsilon^{\mathcal{C}}_{\mathsf{MCA}}(1 - \ell^{-1} + \eta) \geq \frac{1 - 1/|\mathbb{F}|}{\ell - 1} \ . \tag{33}$$

Corollary 9.2 (No uniform exponent above $1/2$)**.** *Fix a prime ℓ and an exponent $\alpha > 1/2$. There exist radii $0 < \gamma < p < 1$ and an integer $L \geq 1$ such that*

$$s := 1 - (1 - p)^\alpha - \gamma > 0$$

and the following holds. For every polynomial $Q \in \mathbb{R}[X, Y, Z]$ and all sufficiently large integers m, there is an $\mathbb{F}_{\ell^m}$-linear, (p, L)-list-decodable code $\mathcal{C} \subseteq \mathbb{F}_{\ell^m}^n$ with $n = O_{\ell,\alpha}(m^2)$ satisfying

$$\epsilon^{\mathcal{C}}_{\mathsf{MCA}}(\gamma) \ \geq \ \frac{1 - \ell^{-m}}{\ell - 1} \ > \ \frac{Q(n, L, s^{-1})}{\ell^m} \ . \tag{34}$$

The radii and list bound are fixed before Q is chosen, so its coefficients may depend on ℓ, α, p, γ without affecting the conclusion.

Comparison with the list-size-dependent bound. The parameters in Theorem 9.1 lie outside the positive-gap range of Theorem 8.1. To see this quantitatively, write $L := L_{\ell,\eta}$, $p := 1 - \ell^{-2} - \eta$, and $\gamma := 1 - \ell^{-1} + \eta$. The choice in (38) gives $L + 1 \geq 4/\eta$, since $L_0 \geq 1$. Hence

$$\begin{aligned}
G &= F_{L+1}(\gamma) - (1 - p) \\
&= -\eta(1 + 2\ell^{-1} - \eta) + \frac{\gamma + (1 - \gamma)\gamma^{L+1}}{L + 1} \\
&\leq -\eta(1 + 2\ell^{-1} - \eta) + \frac{\eta}{4} < 0 \ .
\end{aligned} \tag{35}$$

We used that the numerator of the correction is at most one. Thus the stated list bound, although independent of $|\mathbb{F}|$, does not give a large enough correction to reach the MCA radius of the counterexample. The strict negative margin also persists after rounding the radii for all sufficiently large blocklengths. This explains why the constant MCA error here does not contradict Theorem 8.1.

9.1 A trace subcode

Let $|\mathbb{F}| = \ell^m$ with $m \geq 3$, and evaluate on all of $\mathbb{F}$ the code

$$\mathcal{M}_m := \left\{ z \longmapsto \sum_{j=2}^{m-1} \alpha_j z^{\ell^j} : \alpha_j \in \mathbb{F} \right\} \ . \tag{36}$$

This is an $\mathbb{F}$-linear code of length $|\mathbb{F}|$ and dimension $m-2$, whose words are $\mathbb{F}_\ell$-linear maps $\mathbb{F} \to \mathbb{F}$. Omitting the two lowest Frobenius powers makes z far from the code while allowing many combinations of z and z^ℓ to be close to it.

After the Frobenius change of variables below, the maps have the consecutive linearized-polynomial representation introduced in [Gab85]. Evaluating them on all field elements gives the code in (36), with Hamming distance measured on those evaluations.

Lemma 9.3. *For every $0 < \varepsilon \leq 1 - \ell^{-2}$, the code $\mathcal{M}_m$ is $(1 - \ell^{-2} - \varepsilon, \Lambda_{\ell,\varepsilon})$-list-decodable. The word $\mathbf{u}(z) := z$ satisfies $\delta(\mathbf{u}, \mathcal{M}_m) \geq 1 - \ell^{-2}$.*

Proof. The coordinate permutation $z = \zeta^{\ell^{m-2}}$ turns z^{ℓ^j} into $\zeta^{\ell^{j-2}}$ for $2 \leq j \leq m-1$, using $\zeta^{\ell^m} = \zeta$. Every nonzero codeword therefore becomes a nonzero polynomial of degree at most $\ell^{m-3} = |\mathbb{F}|/\ell^3$, with at most that many roots. Its kernel as an $\mathbb{F}_\ell$-linear map has size at most ℓ^{m-3}, so its rank is at least three. Differences of distinct codewords are nonzero codewords, and Proposition 6.3 gives the list bound.

Under the same permutation, $\mathbf{u}$ becomes $\zeta \mapsto \zeta^{|\mathbb{F}|/\ell^2}$. Its difference from every codeword has degree exactly $|\mathbb{F}|/\ell^2$, and hence at most that many roots. Thus its agreement with any codeword is at most ℓ^{-2}, proving the distance bound. □

The following construction is the prime-field version of the characteristic-two trace example in [BGKS20].

Lemma 9.4. *Let $\mathbf{u}(z) := z$ and $\mathbf{v}(z) := z^\ell$. For $x \in \mathcal{P}$, put*

$$\mathbf{c}_x(z) := -\sum_{j=2}^{m-1} x^{(\ell^j-1)/(\ell-1)} z^{\ell^j} \in \mathcal{M}_m \ .$$

For any $a \in \mathbb{F}^\times$ with $x = a^{\ell-1}$,

$$\mathbf{u}(z) + x\mathbf{v}(z) - \mathbf{c}_x(z) = \frac{\mathsf{Tr}_{\mathbb{F}/\mathbb{F}_\ell}(az)}{a} \ . \tag{37}$$

In particular, $\delta(\mathbf{u} + x\mathbf{v}, \mathbf{c}_x) = 1 - \ell^{-1}$ for every $x \in \mathcal{P}$.

Proof. Expanding the trace gives

$$\begin{aligned} \frac{\mathsf{Tr}_{\mathbb{F}/\mathbb{F}_\ell}(az)}{a} &= \sum_{j=0}^{m-1} a^{\ell^j - 1} z^{\ell^j} \\ &= z + xz^\ell + \sum_{j=2}^{m-1} x^{(\ell^j-1)/(\ell-1)} z^{\ell^j} \ . \end{aligned}$$

The exponents $(\ell^j - 1)/(\ell - 1)$ are integers, so the chosen witness depends only on x, not on the choice of a. The minus sign in its definition gives (37) in every characteristic.

The trace polynomial is nonzero and has degree $|\mathbb{F}|/\ell < |\mathbb{F}|$, so it cannot vanish on the whole field. Thus $z \mapsto \mathsf{Tr}_{\mathbb{F}/\mathbb{F}_\ell}(az)$ is a nonzero $\mathbb{F}_\ell$-linear functional for every $a \neq 0$. Its kernel has $|\mathbb{F}|/\ell$ elements, and scaling by a^{-1} leaves that kernel unchanged. The right-hand side of (37) therefore has relative weight exactly $1 - \ell^{-1}$. □

9.2 Puncturing while preserving list-decodability

The full evaluation code has blocklength $|\mathbb{F}|$, which is too large to obtain the strict inequality in (34) for arbitrary Q. We sample only quadratically many coordinates in $\log |\mathbb{F}|$, preserving the list bound, the distance of $\mathbf{u}$ from the code, and the witnesses for all seeds in $\mathcal{P}$.

Fix $0 < \eta < (\ell^{-1} - \ell^{-2})/2$, and set

$$\begin{aligned} L_0 &:= \Lambda_{\ell,\eta/4} \ , \\ M &:= \left\lceil \frac{4L_0}{\eta} \right\rceil \ , \\ L_{\ell,\eta} &:= M - 1 \ , \\ n &:= \left\lceil \frac{4}{\eta^2} \left(Mm \log |\mathbb{F}| + \log 16 \right) \right\rceil \ . \end{aligned} \tag{38}$$

The ordinary-to-average-radius conversion used below appears in [RW15]; the plurality method for random puncturing is studied in [RW14].

Lemma 9.5 (Coordinatewise multiplicity)**.** *For any distinct $f_1, \ldots, f_M \in \mathcal{M}_m$, the function*

$$Y(z) := \frac{1}{M} \max_{y \in \mathbb{F}} |\{j \in [M] : f_j(z) = y\}|$$

satisfies $\mathbb{E}_{z \in \mathbb{F}}[Y(z)] \leq \ell^{-2} + \eta/2$.

Proof. At each coordinate, choose a most frequent value among the $f_j(z)$, defining a word g. Then

$$\mathbb{E}_{z \in \mathbb{F}}[Y(z)] = \frac{1}{M} \sum_{j=1}^{M} \mathsf{agr}(f_j, g) \ .$$

By Lemma 9.3 with $\varepsilon = \eta/4$, at most L_0 codewords have agreement at least $\ell^{-2} + \eta/4$ with g. Bounding their agreements by one and the others by this threshold gives

$$\mathbb{E}_{z \in \mathbb{F}}[Y(z)] \leq \ell^{-2} + \frac{\eta}{4} + \frac{L_0}{M} \leq \ell^{-2} + \frac{\eta}{2} \ .$$

□

Proof of Theorem 9.1. Choose $z_1, \ldots, z_n \in \mathbb{F}$ independently and uniformly. Let $\mathcal{I} := (z_1, \ldots, z_n)$ and let $f|_{\mathcal{I}} := (f(z_1), \ldots, f(z_n))$. We seek a sample for which all the following estimates hold.

Preserving the list bound. For a fixed list of M distinct original codewords, Lemmas 3.6 and 9.5 give

$$\Pr_{z_1, \ldots, z_n \in \mathbb{F}} \left[\frac{1}{n} \sum_{i=1}^{n} Y(z_i) \geq \ell^{-2} + \eta \right] \leq e^{-n\eta^2/2} \ .$$

There are at most $|\mathbb{F}|^{(m-2)M}$ ordered lists. The probability that any list violates the desired inequality is therefore at most

$$\exp\left(M(m-2)\log|\mathbb{F}| - n\eta^2/2\right) \quad . \tag{39}$$

Outside this event, the restricted code $\mathcal{C}$ is $(1-\ell^{-2}-\eta, M-1)$-list-decodable. Indeed, M distinct punctured words agreeing with a word on at least an $\ell^{-2}+\eta$ fraction would have distinct original representatives. Their average agreement could not exceed their sampled average plurality, contradicting the strict upper bound on that plurality.

Preserving the distances. For each $\mathbf{c} \in \mathcal{M}_m$, the expected relative distance from $\mathbf{u}|_{\mathcal{I}}$ to $\mathbf{c}|_{\mathcal{I}}$ is at least $1-\ell^{-2}$. For each $x \in \mathcal{P}$, the expected relative distance from $(\mathbf{u}+x\mathbf{v})|_{\mathcal{I}}$ to $\mathbf{c}_x|_{\mathcal{I}}$ is $1-\ell^{-1}$. Each distance is an average of independent zero-one variables. Hoeffding's inequality bounds the probability of a downward deviation by more than η in the first case, or an upward deviation by more than η in the second, by $e^{-2\eta^2 n}$. A union bound gives total failure probability at most

$$\left(|\mathbb{F}|^{m-2} + |\mathcal{P}|\right) e^{-2\eta^2 n} \quad . \tag{40}$$

By (38), $n\eta^2/2 \geq 2Mm\log|\mathbb{F}| + 2\log 16$. Thus (39) is at most $1/256$. Also $|\mathcal{P}| \leq |\mathbb{F}| - 1 \leq |\mathbb{F}|^{m-2}$ and $2\eta^2 n \geq 8Mm\log|\mathbb{F}| + 8\log 16$, so (40) is at most $2\cdot 16^{-8}$. Their sum is less than $1/4$.

We also require distinct coordinates. Put $A_\ell := 2\ell^6 + 4$. The length bound from (17) and (38) gives $n \leq C_\ell \eta^{-A_\ell} m^2$ for a constant C_ℓ. Hence the probability of a collision is at most

$$\frac{n(n-1)}{2\ell^m} \leq \frac{C_\ell^2 \eta^{-2A_\ell} m^4}{2\ell^m} \quad .$$

This is at most $1/4$ for every $m \geq m_0$, where $m_0 = O_\ell(\log(2/\eta))$: after taking logarithms, the required inequality is $m\log\ell \geq 2A_\ell\log(1/\eta) + 4\log m + O_\ell(1)$, and its linear left-hand side dominates the logarithmic term. Increase m_0 to at least three. The total failure probability is then less than one, so a sample preserving all estimates and having distinct coordinates exists for every $m \geq m_0$.

Denote the restricted code and words by $\mathcal{C}, \mathbf{u}, \mathbf{v}$. The bounds on $L_{\ell,\eta}$ and n in (31) follow from (17) and (38):

$$\begin{aligned} L_0 &= O_\ell\left(\eta^{-(2\ell^6+1)}\right) \quad , \\ L_{\ell,\eta} &= O_\ell\left(\eta^{-(2\ell^6+2)}\right) \quad , \\ n &= O_\ell\left(\eta^{-(2\ell^6+4)} m^2\right) \quad . \end{aligned}$$

All required distance and list-decoding estimates now hold simultaneously.

Witnessing MCA-badness. Fix $x \in \mathcal{P}$, and write $\mathbf{c}_x$ for its punctured witness. Let $A(x)$ be the agreement set of $\mathbf{u}+x\mathbf{v}$ and $\mathbf{c}_x$. Its size is at least $(\ell^{-1}-\eta)n$. If $\mathbf{u}|_{A(x)} \in \mathcal{C}|_{A(x)}$, a codeword agreeing with $\mathbf{u}$ on that set would give

$$\delta(\mathbf{u}, \mathcal{C}) \leq 1 - \frac{|A(x)|}{n} \leq 1 - \ell^{-1} + \eta \quad .$$

This contradicts $\delta(\mathbf{u}, \mathcal{C}) \geq 1 - \ell^{-2} - \eta$, because $2\eta < \ell^{-1} - \ell^{-2}$. Hence no codeword agrees with $\mathbf{u}$ on the set where the combination agrees with $\mathbf{c}_x$. It witnesses that x is MCA-bad at radius $1 - \ell^{-1} + \eta$.

Every seed in $\mathcal{P}$ is therefore bad. By (32), the MCA error is at least $|\mathcal{P}|/|\mathbb{F}| = (1 - 1/|\mathbb{F}|)/(\ell - 1)$, proving (33). □

9.3 Proof of Corollary 9.2

Proof. Fix ℓ and $\alpha > 1/2$, and put $\bar{\alpha} := \min\{\alpha, 1\}$ and $\theta := \bar{\alpha} - 1/2$. Choose

$$D := \ell^{-1} - \ell^{-2\bar{\alpha}} > 0, \qquad \eta := \frac{D}{4(\ell+1)} \ .$$

Then $0 < \eta < (\ell^{-1} - \ell^{-2})/2$. On $x \geq \ell^{-2}$, the derivative of $x^{\bar{\alpha}}$ is at most ℓ, so, with

$$p := 1 - \ell^{-2} - \eta \ , \qquad \gamma := 1 - \ell^{-1} + \eta \ ,$$

the slack

$$s = 1 - (1-p)^{\alpha} - \gamma = \ell^{-1} - \eta - (\ell^{-2} + \eta)^{\alpha}$$

satisfies

$$s \geq \ell^{-1} - \eta - (\ell^{-2} + \eta)^{\bar{\alpha}} \geq D - (\ell+1)\eta = \frac{3D}{4} > 0 \ .$$

These radii also satisfy $0 < \gamma < p < 1$ by the choice of η. Moreover,

$$D = \ell^{-1}(1 - \ell^{-2\theta}) = \Theta_{\ell}(\theta), \qquad \eta = \Theta_{\ell}(\theta) \ .$$

Thus Theorem 9.1 constructs the codes for every $m \geq m_0 = O_{\ell}(\log(2/\theta))$, with the list and length bounds stated in Theorem 3. Set $L := L_{\ell,\eta}$ and let $|\mathbb{F}| = \ell^m \to \infty$. The parameters p, γ, s, L depend only on ℓ, α and stay fixed, while $n = O_{\ell,\alpha}(m^2)$. For every fixed polynomial Q,

$$\frac{Q(n, L_{\ell,\eta}, s^{-1})}{|\mathbb{F}|} \longrightarrow 0 \ ,$$

whereas

$$\epsilon^{\mathcal{C}}_{\mathsf{MCA}}(\gamma) \geq \frac{1 - 1/|\mathbb{F}|}{\ell - 1} \longrightarrow \frac{1}{\ell - 1} > 0 \ .$$

Thus the strict inequality in (34) holds for all sufficiently large m. Only this final comparison requires a threshold depending on Q. □

10 Codes with large list size and small MCA error

We use $q = \ell^m$ for the size of the base field $\mathbb{F}_q$, where ℓ is any prime and m is a positive integer. For a finite extension $\mathbb{K}/\mathbb{F}_q$ of degree e, write $|\mathbb{K}| = q^e$. Recall the notation $J_2(d) := 1 - \sqrt{1-d}$. In this section we also use:

$$\Psi(\tau) := 1 - (1-\tau)^3 \ . \tag{41}$$

The function Ψ is increasing and concave on $[0,1]$.

Theorem 10.1. *Fix a prime ℓ and*

$$\begin{aligned} &0 < d < 1 \ , \\ &J_2(d) < p < \gamma < 1 \ . \end{aligned} \tag{42}$$

Put

$$\xi := \min\{d, 1-d, p - J_2(d), (1-\gamma)^3\} > 0 \ . \tag{43}$$

There exist a power q of ℓ, integers $k \geq 2$ and $n_0 \geq 1$, constants $R, c_0, C_0, \beta > 0$, and linear codes $\mathcal{C}_n \subseteq \mathbb{F}_q^n$ at an unbounded sequence of lengths n, such that every scalar extension

$$\begin{aligned} &\mathcal{C}_{n,\mathbb{K}} := \mathcal{C}_n \otimes_{\mathbb{F}_q} \mathbb{K} \ , \\ &\mathbb{K}/\mathbb{F}_q \text{ a finite extension of degree } e \geq 1 \ , \\ &|\mathbb{K}| = q^e \ , \end{aligned}$$

satisfies

$$\begin{aligned} &\frac{\dim_{\mathbb{K}} \mathcal{C}_{n,\mathbb{K}}}{n} \longrightarrow R \ , \\ &\delta(\mathcal{C}_{n,\mathbb{K}}) \longrightarrow d \ , \end{aligned} \tag{44}$$

and, for every length $n \geq n_0$ in the sequence,

$$\begin{aligned} &L_{\mathcal{C}_{n,\mathbb{K}}}(p) \ \geq \ 2^{c_0 n^{\beta} \log_2 n} \ , \\ &\epsilon_{\mathsf{MCA}}^{\mathcal{C}_{n,\mathbb{K}}}(\gamma) \ \leq \ \frac{C_0 n^k}{|\mathbb{K}|} \ . \end{aligned} \tag{45}$$

All constants depend only on ℓ and the fixed parameters in (42)*, not on n or $|\mathbb{K}|$. The parameters can be chosen with*

$$\begin{aligned} &q = O_\ell(\xi^{-2}), \qquad \beta^{-1} = O_\ell\left(\xi^{-1} \log \frac{2}{\xi}\right), \\ &k, n_0 \leq \exp\left(\exp\left(O_\ell\left(\xi^{-1} \log^2 \frac{2}{\xi}\right)\right)\right) \ . \end{aligned} \tag{46}$$

Note that, for every such length, taking $e = \lceil (2k+5) \log_q n \rceil$ gives

$$n^{2k+5} \leq |\mathbb{K}| \leq n^{2k+6}, \qquad \epsilon_{\mathsf{MCA}}^{\mathcal{C}_{n,\mathbb{K}}}(\gamma) \leq n^{-2} \ .$$

Corollary 10.2 (No polynomial strong converse)**.** *Fix a prime ℓ, $0 < d < 1$, and radii $J_2(d) < p < \gamma < 1$. Put $s := \gamma - p > 0$. There exist constants $R, A > 0$, an integer n_0, and a family of linear codes $\mathcal{D}_n \subseteq \mathbb{K}_n^n$ at an unbounded sequence of lengths n, whose rates tend to R and whose relative distances tend to d, with the following properties. Each $\mathbb{K}_n$ has characteristic ℓ, and, for every length $n \geq n_0$ in the sequence,*

$$n \leq |\mathbb{K}_n| \leq n^A, \qquad \epsilon_{\mathsf{MCA}}^{\mathcal{D}_n}(\gamma) \leq n^{-1} \ .$$

One may take $A = 2k + 6$, with k and n_0 satisfying (46)*. For every polynomial $Q \in \mathbb{R}[X, Y, Z]$, the same family satisfies*

$$L_{\mathcal{D}_n}(p) \ > \ Q\big(n, |\mathbb{K}_n| \epsilon_{\mathsf{MCA}}^{\mathcal{D}_n}(\gamma), s^{-1}\big) \tag{47}$$

for all sufficiently large n. The family and the constants R, A depend only on ℓ, d, p, γ.

Thus no polynomial in the blocklength, normalized MCA error, and inverse radius gap bounds the list size in this regime. The parameters are fixed before Q is chosen, so its coefficients may depend on them. The limiting rate R is positive but is not independently prescribed.

10.1 Generalized weights and mutual correlated agreement

Generalized Hamming weights were introduced in [Wei91].

Definition 10.3 (Shortened codes and generalized weights)**.** *For a linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$ and $U \subseteq [n]$ define the shortened code over U to be $\mathcal{C}[U] := \{\mathbf{b} \in \mathcal{C} : \mathsf{supp}(\mathbf{b}) \subseteq U\}$. For $1 \leq j \leq \dim \mathcal{C}$ the j-th generalized Hamming weight of $\mathcal{C}$ is*

$$d_j(\mathcal{C}) := \min_{\substack{V \leq \mathcal{C} \\ \dim V = j}} |\mathsf{supp}(V)| = \min\{|U| \ : \ \dim \mathcal{C}[U] \geq j\} \ ,$$
$$\mathsf{supp}(V) := \bigcup_{\mathbf{v} \in V} \mathsf{supp}(\mathbf{v}) \ .$$

The two descriptions agree: if $\dim \mathcal{C}[U] \geq j$, any j-dimensional $V \leq \mathcal{C}[U]$ has $\mathsf{supp}(V) \subseteq U$; conversely $V \leq \mathcal{C}[\mathsf{supp}(V)]$.

Lemma 10.4 (Scalar extension)**.** *Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be a linear code, let $\mathbb{K}/\mathbb{F}_q$ be a finite extension, and put $\mathcal{C}_{\mathbb{K}} := \mathcal{C} \otimes_{\mathbb{F}_q} \mathbb{K} \subseteq \mathbb{K}^n$. For every $U \subseteq [n]$,*

$$\mathcal{C}_{\mathbb{K}}[U] = \mathcal{C}[U] \otimes_{\mathbb{F}_q} \mathbb{K} \ ,$$
$$\dim_{\mathbb{K}}(\mathcal{C}_{\mathbb{K}}[U]) = \dim_{\mathbb{F}_q}(\mathcal{C}[U]) \ .$$

Consequently, scalar extension preserves the code dimension and all generalized Hamming weights:

$$\dim_{\mathbb{K}} \mathcal{C}_{\mathbb{K}} = \dim_{\mathbb{F}_q} \mathcal{C} \ ,$$
$$d_j(\mathcal{C}_{\mathbb{K}}) = d_j(\mathcal{C}) \qquad \big(1 \leq j \leq \dim_{\mathbb{F}_q} \mathcal{C}\big) \ .$$

Proof. Identify $\mathcal{C}_{\mathbb{K}}$ with the $\mathbb{K}$-linear span of $\mathcal{C}$ in $\mathbb{K}^n$, and fix $U \subseteq [n]$. We first show that shortening to U commutes with this extension, and then compare the dimensions over the respective fields.

Fix an $\mathbb{F}_q$-basis $\xi_1, \ldots, \xi_e$ of $\mathbb{K}$. Every $\mathbf{c} \in \mathcal{C}_{\mathbb{K}}$ has a unique expansion

$$\mathbf{c} = \sum_{l=1}^{e} \xi_l \mathbf{b}_l \ , \qquad \mathbf{b}_l \in \mathcal{C} \ .$$

To obtain this expansion, write $\mathbf{c}$ as a $\mathbb{K}$-linear combination of words in $\mathcal{C}$ and expand each coefficient in the basis $\xi_1, \ldots, \xi_e$. The resulting component words $\mathbf{b}_l$ lie in $\mathcal{C}$ by $\mathbb{F}_q$-linearity. Uniqueness follows coordinatewise from the linear independence of the ξ_l over $\mathbb{F}_q$. For the same reason, $\mathbf{c}[i] = 0$ if and only if $\mathbf{b}_l[i] = 0$ for every l. Thus $\mathbf{c}$ is supported on U if and only if all its component words lie in $\mathcal{C}[U]$. Conversely, every $\mathbb{K}$-linear combination of words in $\mathcal{C}[U]$ is supported on U. This proves

$$\mathcal{C}_{\mathbb{K}}[U] = \mathcal{C}[U] \otimes_{\mathbb{F}_q} \mathbb{K} \ .$$

To see why the dimensions agree, let $m := \dim_{\mathbb{F}_q}(\mathcal{C}[U])$ and choose an $\mathbb{F}_q$-basis $\mathbf{v}_1, \ldots, \mathbf{v}_m$ of $\mathcal{C}[U]$. The identity just proved shows that these same vectors span $\mathcal{C}_{\mathbb{K}}[U]$ over $\mathbb{K}$. They also remain linearly independent over $\mathbb{K}$. Indeed, suppose that $\sum_{j=1}^{m} a_j \mathbf{v}_j = 0$ with $a_j \in \mathbb{K}$, and write $a_j = \sum_{l=1}^{e} \xi_l a_{l,j}$ with $a_{l,j} \in \mathbb{F}_q$. Then

$$0 = \sum_{l=1}^{e} \xi_l \left(\sum_{j=1}^{m} a_{l,j} \mathbf{v}_j \right) \ .$$

Independence of the ξ_l, applied in each coordinate, gives $\sum_j a_{l,j} \mathbf{v}_j = 0$ for every l. Independence of the $\mathbf{v}_j$ over $\mathbb{F}_q$ then forces every $a_{l,j}$, and hence every a_j, to be zero. The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ therefore form a $\mathbb{K}$-basis of the extended shortened code, proving

$$\dim_{\mathbb{K}}(\mathcal{C}_{\mathbb{K}}[U]) = \dim_{\mathbb{K}}\big(\mathcal{C}[U] \otimes_{\mathbb{F}_q} \mathbb{K}\big) = \dim_{\mathbb{F}_q}(\mathcal{C}[U]) \ .$$

Taking $U = [n]$ gives preservation of the code dimension. For each j, the condition that the shortened code on U have dimension at least j is also unchanged. Minimizing $|U|$ subject to that condition, as in Definition 10.3, proves preservation of every generalized Hamming weight. □

The next lemma bounds the MCA error using a single generalized Hamming weight.

Lemma 10.5 (MCA from a generalized weight)**.** *Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be linear, let $1 \le k \le \dim \mathcal{C}$, and let $\mathcal{C}_{\mathbb{K}} = \mathcal{C} \otimes_{\mathbb{F}_q} \mathbb{K}$. Put*

$$\mathsf{T}_k(n, q) := \sum_{j=1}^{k} \binom{n}{j} (q-1)^{j-1} \ . \tag{48}$$

If $d_k(\mathcal{C}) > n\Psi(\gamma)$, then

$$|\mathbb{K}|\, \epsilon_{\mathsf{MCA}}^{\mathcal{C}_{\mathbb{K}}}(\gamma) \ \le \ \frac{\mathsf{T}_k(n, q) + 3}{1 - n\Psi(\gamma)/d_k(\mathcal{C})} \ . \tag{49}$$

Proof. For each fixed pair of distinct bad seeds, we bound the number of third seeds whose three error sets have union size below $d_k(\mathcal{C})$. We associate each such third seed with an equation involving at most k coordinates and coefficients in $\mathbb{F}_q$. Each equation determines at most one third seed.

Fix $\mathbf{u}, \mathbf{v} \in \mathbb{K}^n$, and let X be their set of bad seeds, of size M. The case $M = 0$ is immediate. For each $x \in X$, choose a witness agreement set T_x for the failure of MCA, put $E_x = [n] \setminus T_x$, and choose $\mathbf{c}_x \in \mathcal{C}_{\mathbb{K}}$ such that

$$\begin{aligned} &|E_x| \le \gamma n \ , \\ &\mathsf{supp}(\mathbf{u} + x\mathbf{v} - \mathbf{c}_x) \subseteq E_x \ . \end{aligned}$$

Write $d_* = d_k(\mathcal{C})$.

The k-th generalized weight bounds shortening dimensions. From Lemma 10.4, scalar extension commutes with shortening:

$$\mathcal{C}_{\mathbb{K}}[U] = \mathcal{C}[U] \otimes_{\mathbb{F}_q} \mathbb{K} \ .$$

Consequently,

$$\begin{aligned} &|U| < d_* \\ &\implies \dim_{\mathbb{K}} \mathcal{C}_{\mathbb{K}}[U] \leq k-1 \ . \end{aligned} \tag{50}$$

Thus codewords known to vanish outside U have at most $k-1$ degrees of freedom.

Two fixed seeds localize the problem. Fix distinct $x_0, x_1 \in X$. Let $Q_0(\mathtt{X})$ be the unique affine polynomial with coefficients in $\mathcal{C}_{\mathbb{K}}$ satisfying $Q_0(x_0) = \mathbf{c}_{x_0}$ and $Q_0(x_1) = \mathbf{c}_{x_1}$, and write

$$\begin{aligned} f(\mathtt{X}) &= \mathbf{u} + \mathtt{X}\mathbf{v} - Q_0(\mathtt{X}) = \mathbf{a} + \mathtt{X}\mathbf{b} \ , \\ S &= E_{x_0} \cup E_{x_1} \ . \end{aligned}$$

Outside S, f vanishes at both x_0 and x_1, and is therefore identically 0. Hence $\mathsf{supp}(\mathbf{a}), \mathsf{supp}(\mathbf{b}) \subseteq S$. Write $Q_0(\mathtt{X}) = \mathbf{q}_0 + \mathtt{X}\mathbf{q}_1$ with $\mathbf{q}_0, \mathbf{q}_1 \in \mathcal{C}_{\mathbb{K}}$. For every $x \in X$, the word $\mathbf{a} + x\mathbf{b}$ equals $\mathbf{c}_x - Q_0(x)$ on T_x. Also, $\mathbf{a} = \mathbf{u} - \mathbf{q}_0$ equals a codeword on T_x if and only if $\mathbf{u}$ does, and likewise for $\mathbf{b} = \mathbf{v} - \mathbf{q}_1$ and $\mathbf{v}$. Thus at least one of $\mathbf{a}, \mathbf{b}$ does not equal any codeword on T_x. Consider a third seed $x_2 \in X \setminus \{x_0, x_1\}$ such that

$$|E_{x_0} \cup E_{x_1} \cup E_{x_2}| < d_* \ .$$

Set

$$\begin{aligned} U &= S \cup E_{x_2} \ , \\ J &= S \cap T_{x_2} = S \setminus E_{x_2} \ , \\ W &= \mathcal{C}_{\mathbb{K}}[U]|_J \ . \end{aligned}$$

The codeword $\mathbf{c}_{x_2} - Q_0(x_2)$ equals $f(x_2)$ on T_{x_2}. Since $f(x_2)$ is supported on S, this codeword vanishes on $T_{x_2} \setminus S = [n] \setminus U$ and is therefore supported on U. More generally, a word g supported on S equals a codeword on T_{x_2} if and only if that codeword is supported on U and restricts to g on J. By definition of W, this is equivalent to $g|_J \in W$. Applying this equivalence to $\mathbf{a} + x_2\mathbf{b}$, $\mathbf{a}$ and $\mathbf{b}$ gives

$$\begin{aligned} &(\mathbf{a} + x_2\mathbf{b})|_J \in W \ , \\ &(\mathbf{a}|_J, \mathbf{b}|_J) \notin W^2 \ . \end{aligned} \tag{51}$$

By (50), $\dim W \leq k-1$.

Expressing membership by equations on at most k coordinates. The space W is defined over $\mathbb{F}_q$, and since scalar extension preserves dimensions, we can choose an information set $I \subseteq J$ of size $r = \dim W$. Membership in W is expressed by equations

$$\begin{aligned} &w_j - \sum_{i \in I} \alpha_{j,i} w_i = 0 \\ &(j \in J \setminus I) \ , \\ &\alpha_{j,i} \in \mathbb{F}_q \ . \end{aligned}$$

Each equation involves at most $r+1 \le k$ coordinates. By (51), some such equation is satisfied by $(\mathbf{a}+x_2\mathbf{b})|_J$ but is not satisfied by both $\mathbf{a}|_J$ and $\mathbf{b}|_J$. Extending its coefficient vector by zero outside J gives $\lambda \in \mathbb{F}_q^n$ with

$$\begin{aligned} &|\mathsf{supp}(\lambda)| \le k \ , \\ &\lambda \cdot \mathbf{a} + x_2 \lambda \cdot \mathbf{b} = 0 \ , \\ &(\lambda \cdot \mathbf{a}, \lambda \cdot \mathbf{b}) \ne (0,0) \ . \end{aligned}$$

Thus x_2 is a root of the nonzero affine polynomial $\lambda \cdot \mathbf{a} + \mathtt{X}\lambda \cdot \mathbf{b}$. Normalize the first nonzero entry of λ to one. There are exactly $\mathsf{T}_k(n,q)$ possible normalized vectors of support size at most k. For fixed x_0, x_1, the words $\mathbf{a}, \mathbf{b}$ are fixed, and each vector gives at most one root. Hence at most $\mathsf{T}_k(n,q)$ third seeds have union of error sets smaller than d_*. Although U, J, W, and their information sets can depend on the third seed, every resulting coefficient vector belongs to this same collection of $\mathsf{T}_k(n,q)$ vectors.

Counting triples with union size below d_*. We use the expected union-of-disagreements computation from DEEP-FRI [BGKS20]. Draw x_0, x_1, x_2 independently and uniformly from X, and set $p_i = \Pr_{x \in X}[i \in E_x]$. Independence gives

$$\mathbb{E}\big[|E_{x_0} \cup E_{x_1} \cup E_{x_2}|\big] = \sum_{i=1}^{n} \big(1 - (1-p_i)^3\big) \ .$$

Since $\sum_i p_i \le \gamma n$ and Ψ is increasing and concave on $[0,1]$, Jensen's inequality yields

$$\mathbb{E}\big[|E_{x_0} \cup E_{x_1} \cup E_{x_2}|\big] \le n\Psi\left(\frac{1}{n}\sum_i p_i\right) \le n\Psi(\gamma) \ .$$

Markov's inequality implies that at least $\left(1 - \frac{n\Psi(\gamma)}{d_k(\mathcal{C})}\right) M^3$ ordered triples have union size less than d_*. At most $3M^2$ ordered triples have a repeated seed. Counting the remaining triples by their ordered pair (x_0, x_1) therefore gives

$$\left(1 - \frac{n\Psi(\gamma)}{d_k(\mathcal{C})}\right) M^3 - 3M^2 \le M(M-1)\mathsf{T}_k(n,q) \le M^2 \mathsf{T}_k(n,q) \ .$$

Dividing by M^2 and rearranging gives

$$M \le \frac{\mathsf{T}_k(n,q) + 3}{1 - n\Psi(\gamma)/d_k(\mathcal{C})} \ ,$$

as required. □

Relation to earlier arguments. The support–dimension interpretation of generalized weights is standard. It is used in [GGR09] to reduce list-decoding of interleaved and tensor-product codes to low-rank cases. Dimension bounds and coordinate sampling also underlie the MCA results for subspace-design codes in [GG26]; [AGG26] extends this approach from subfield subspace-design bounds to extension-field line decoding.

A closer precedent is [Cho26b]. For a fixed coordinate set U, a parity-check matrix H over $\mathbb{K}$, and $r = \dim_{\mathbb{K}} \ker H_U$, it bounds the number of exceptional slopes with transverse witnesses supported on at most t coordinates of U by $(t+1)|\mathbb{K}|^r$. Here transversality means that the combined word agrees with a codeword on the witness set, while at least one of the two original words does not agree with any codeword on that set. The same work gives a bound independent of the field size for a fixed MDS circuit. A public

companion draft also uses generalized weights to bound lists inside affine codeword spaces [Cho26a]. In the rank metric, [LXYZ26] subtracts two witnesses and counts base-field linear dependencies to bound exceptional parameters; minimum rank distance forces its residual codeword to vanish.

The contribution of Lemma 10.5 is to use one higher generalized Hamming weight to allow positive-dimensional shortenings, while counting equations over the fixed base field uniformly as U varies. This gives a numerator independent of $|\mathbb{K}|$ for every scalar extension, without a list-size hypothesis. The three-witness averaging step comes from DEEP-FRI [BGKS20]; the additional step is the short-equation count that replaces uniqueness of the localized codeword.

10.2 Symmetric forms with large lists and large supports

We start with symmetric multilinear forms over $\mathbb{F}_\ell$, evaluated as vector-valued words. Each vector-valued symbol is treated as a single block. The relative block distance between two words is the fraction of blocks on which their symbols differ.

We exhibit a word within a small relative block distance of many distinct codewords. Symmetry gives lower bounds on the code's minimum block distance and on the support of every sufficiently high-dimensional subspace. We then amplify these supports to obtain scalar-valued codes of fixed positive rate and the prescribed relative distance. This allows us to use our scalar-extension and MCA results in Lemmas 10.4 and 10.5. A simpler construction over binary fields using BCH codes is given in Appendix C.

Symmetric forms and their supports. Let $H := \mathbb{F}_\ell^h$, with standard basis $e_1, \dots, e_h$, and let $r \geq 1$ be an integer.

Definition 10.6 (Multilinear and symmetric forms)**.** *An r-linear form, or multilinear form, on H is a map $T : H^r \to \mathbb{F}_\ell$ that is linear in each argument when the others are fixed. It is symmetric if*

$$T(z_1, \dots, z_r) = T(z_{\sigma(1)}, \dots, z_{\sigma(r)})$$

for every $z_1, \dots, z_r \in H$ and every permutation σ of $[r]$. Write $\mathcal{M}_r(H)$ for the $\mathbb{F}_\ell$-space of all r-linear forms on H, and $\mathcal{S}_r(H)$ for its subspace of symmetric forms. Their dimensions are

$$\dim \mathcal{M}_r(H) = h^r \ , \qquad \dim \mathcal{S}_r(H) = \binom{h+r-1}{r} \ .$$

A multilinear form is determined by its values on the h^r ordered tuples of standard basis vectors, and these values can be chosen arbitrarily. For a symmetric form, each value depends only on the multiplicities of the indices, giving one independent value for each $(\alpha_1, \dots, \alpha_h) \in \mathbb{N}^h$ with $\sum_i \alpha_i = r$.

Fix an integer $t \geq 1$.

Definition 10.7 (Evaluation words and block support)**.** *For $T \in \mathcal{S}_{t+1}(H)$, define its evaluation word $\Phi_T : H^t \to H^\vee$ by*

$$\Phi_T(x_1, \dots, x_t) := T(x_1, \dots, x_t, \cdot) \ .$$

Each position of this word is one $H^\vee$-valued block, indexed by a tuple in H^t. The map $T \mapsto \Phi_T$ is linear and injective: if $\Phi_T = 0$, then $T(x_1, \dots, x_t, z) = 0$ for all $x_1, \dots, x_t, z \in H$, so $T = 0$.

For a subspace $V \leq \mathcal{S}_{t+1}(H)$, define its relative block support by

$$\mu_t(V) := \Pr_{x_1,\dots,x_t \in H}[\Phi_T(x_1, \dots, x_t) \neq 0 \text{ for some } T \in V] \ ,$$

where the arguments are independent and uniform. Thus $\mu_t(V)$ is the fraction of blocks on which at least one evaluation word from V is nonzero. For an individual form, write $\mu_t(T) := \mu_t(\langle T \rangle)$; this is the relative block weight of Φ_T.

A nonzero form is a pure power if $T = c\, v^{\otimes(t+1)}$ for some $c \in \mathbb{F}_\ell^\times$ and $v \in H^\vee \setminus \{0\}$, where

$$v^{\otimes(t+1)}(z_1, \ldots, z_{t+1}) := \prod_{j=1}^{t+1} v(z_j) \ .$$

Set

$$\begin{aligned} a_\ell &:= 1 - 1/\ell \ , \\ \beta_t &:= a_\ell^t \ , \\ \delta_t &:= 2\beta_t - \beta_t^2 \ , \\ G_t &:= \beta_t \left(1 + \frac{t}{\ell - 1}\right) \ . \end{aligned} \tag{52}$$

Lemma 10.8 (Minimum support)**.** *Every nonzero $T \in \mathcal{S}_{t+1}(H)$ satisfies $\mu_t(T) \geq \beta_t$. If T is not a pure power, then $\mu_t(T) \geq \delta_t$.*

Proof. For $x \in H$, fix the first argument of T at x and write $T_x := T(x, \cdot, \ldots, \cdot) \in \mathcal{S}_t(H)$. If $T \neq 0$, the map $x \mapsto T_x$ is a nonzero linear map, so its kernel has codimension at least one. Thus $T_x \neq 0$ with probability at least a_ℓ for a uniform $x \in H$. Repeating this argument as we fix the first t arguments successively gives $\mu_t(T) \geq a_\ell^t = \beta_t$. A nonzero pure power attains equality.

For a nonzero symmetric form of order at least two, the map obtained by fixing its first argument has rank one if and only if the form is a pure power. To see this for T, suppose that $x \mapsto T_x$ has rank one and choose a nonzero form Q spanning its image. For every $x \in H$, there is then a unique scalar $v(x) \in \mathbb{F}_\ell$ such that

$$T_x = v(x) Q \ .$$

Since $x \mapsto T_x$ is linear and $Q \neq 0$, these scalars define a nonzero linear functional $v \in H^\vee$. Moreover, $T_x = 0$ if and only if $v(x) = 0$, so the kernel of $x \mapsto T_x$ is exactly $\ker v$. Thus T vanishes whenever its first argument lies in $\ker v$, and symmetry gives the same conclusion for every argument. It therefore factors through the one-dimensional quotient $H/\ker v$ in every argument, and is a nonzero scalar multiple of $v^{\otimes(t+1)}$. Conversely, for a pure power $T = c\, v^{\otimes(t+1)}$, we have $T_x = c\, v(x) v^{\otimes t}$, so $x \mapsto T_x$ has rank one.

For the second assertion, the case $t = 1$ is the matrix rank bound: a symmetric bilinear form that is not a pure power has rank at least two, and hence

$$\mu_1(T) \geq 1 - \ell^{-2} = \delta_1 \ .$$

Suppose $t \geq 2$ and the assertion holds for $t - 1$. Consider the linear map

$$A : H \longrightarrow \mathcal{S}_t(H) \ , \qquad A(x) := T_x \ .$$

Since T is nonzero and not a pure power, the rank-one characterization above gives $\mathsf{rank}(A) \geq 2$. Its kernel therefore has codimension at least two in H. For a uniformly chosen $x \in H$, the probability that T_x is zero is consequently

$$z_0 := \Pr_{x \in H}[T_x = 0] = \frac{|\ker A|}{|H|} = \ell^{-\mathsf{rank}(A)} \leq \ell^{-2} \ .$$

We next need to control how often T_x is a nonzero pure power. This requires examining the rank of a second map. For each $Q \in \mathcal{S}_t(H)$, put

$$B_Q : H \longrightarrow \mathcal{S}_{t-1}(H) \ , \qquad B_Q(y) := Q(y, \cdot, \ldots, \cdot) \ .$$

The same rank-one characterization applies to these t-linear forms, since $t \geq 2$: for $Q \neq 0$, the map B_Q has rank one exactly when Q is a pure power. We first show that the image $\mathcal{W} := A(H)$ contains a form Q with $\mathsf{rank}(B_Q) \geq 2$.

Since $\dim \mathcal{W} = \mathsf{rank}(A) \geq 2$, choose linearly independent $Q_1, Q_2 \in \mathcal{W}$. If either associated map already has rank at least two, we are done. Otherwise both maps have rank one, so

$$Q_i = c_i \, v_i^{\otimes t} \ , \qquad c_i \in \mathbb{F}_\ell^\times \ , \quad v_i \in H^\vee \setminus \{0\} \qquad (i = 1, 2) \ .$$

The functionals v_1, v_2 are linearly independent, since proportional functionals would give proportional forms. Thus the map $y \mapsto (v_1(y), v_2(y))$ is onto $\mathbb{F}_\ell^2$. Choose $y_1, y_2 \in H$ with

$$(v_1(y_1), v_2(y_1)) = (1, 0) \ , \qquad (v_1(y_2), v_2(y_2)) = (0, 1) \ .$$

For $Q := Q_1 + Q_2 \in \mathcal{W}$, we obtain

$$B_Q(y_1) = c_1 v_1^{\otimes(t-1)} \ , \qquad B_Q(y_2) = c_2 v_2^{\otimes(t-1)} \ .$$

These two images are linearly independent: evaluating the forms at $(y_1, \ldots, y_1)$ and $(y_2, \ldots, y_2)$ isolates their respective coefficients. Here we use $t - 1 \geq 1$. Hence $\mathsf{rank}(B_Q) \geq 2$, as claimed. Since $Q \in A(H)$, there is an $x_* \in H$ with $T_{x_*} = Q$ and $\mathsf{rank}(B_{T_{x_*}}) \geq 2$.

We now express B_{T_x} as a matrix while allowing x to vary. Use the standard basis $e_1, \ldots, e_h$ in its domain H. In its codomain $\mathcal{S}_{t-1}(H)$, represent a form by its values on the basis tuples $(e_{i_2}, \ldots, e_{i_t})$ with $1 \leq i_2 \leq \cdots \leq i_t \leq h$. These values determine the form and can be chosen independently. Thus the matrix $M(x)$ has columns indexed by $j \in [h]$ and rows indexed by these tuples, with entries

$$\begin{aligned} M(x)_{(i_2,\ldots,i_t),j} &= T(x, e_j, e_{i_2}, \ldots, e_{i_t}) \\ &= \sum_{i=1}^{h} x_i \, T(e_i, e_j, e_{i_2}, \ldots, e_{i_t}) \ . \end{aligned}$$

Each entry is therefore a homogeneous linear polynomial in the coordinates of x; the values of T on the right are fixed scalars. At the point x_* found above, this matrix has rank at least two, so some 2×2 minor has nonzero determinant. Fix those two rows and columns, and let $P(x)$ be the determinant of that submatrix of $M(x)$. Each term of the determinant is a product of two homogeneous linear polynomials, and $P(x_*) \neq 0$. Hence P is a nonzero homogeneous quadratic polynomial.

If $T_x = 0$, then $M(x) = 0$. If T_x is a nonzero pure power, then $M(x)$ has rank one. In either case all its 2×2 minors vanish, so $P(x) = 0$. A nonzero homogeneous quadratic on $\mathbb{F}_\ell^h$ vanishes on at most a fraction $2/\ell - 1/\ell^2$ of the points, including when $\ell = 2$. This is the degree-two case of the projective hypersurface bound; see [GL24]. Consequently,

$$\begin{aligned} s_0 &:= \Pr_{x \in H}[T_x = 0 \text{ or } T_x \text{ is a pure power}] \\ &\leq \Pr_{x \in H}[P(x) = 0] \leq s_* := 2/\ell - 1/\ell^2 \ . \end{aligned}$$

Put $b := a_\ell^{t-1}$. Nonzero pure powers among the forms T_x have relative block weight b, and the induction hypothesis gives weight at least $2b - b^2$ for every other nonzero T_x. Conditioning on the first argument gives $\mu_t(T) = \mathbb{E}_{x \in H}[\mu_{t-1}(T_x)]$. The probabilities of the zero, nonzero pure-power, and remaining cases are z_0, $s_0 - z_0$, and $1 - s_0$, respectively. Since $2b - b^2 \geq b$, we obtain

$$\begin{aligned}\mu_t(T) &\geq (s_0 - z_0)b + (1 - s_0)(2b - b^2) \\ &\geq (s_* - \ell^{-2})b + (1 - s_*)(2b - b^2) \\ &= 2a_\ell b - a_\ell^2 b^2 = \delta_t \ .\end{aligned}$$

□

To obtain a large generalized weight, we must bound the joint support of entire spaces of forms. Fixing the first argument lets us reduce the order of the forms and argue inductively. For $V \leq \mathcal{S}_{t+1}(H)$ and $x \in H$, put

$$V_x := \{T_x : T \in V\} \leq \mathcal{S}_t(H) \ .$$

The next lemma ensures that, when V has sufficiently large dimension, V_x still has large dimension for many choices of x. The dimension threshold is independent of h; this will allow us to bound a fixed generalized weight as h grows.

Lemma 10.9 (Dimensions after fixing one argument)**.** *Fix $t \geq 1$, an integer $s \geq 1$, and $\varepsilon > 0$. There is an integer K, depending only on t, ℓ, s, ε, such that, for every h and every $V \leq \mathcal{S}_{t+1}(\mathbb{F}_\ell^h)$ with $\dim V \geq K$,*

$$\Pr_{x \in \mathbb{F}_\ell^h}[\dim V_x \geq s] \geq a_\ell - \varepsilon \ .$$

Proof. We first prove that a uniform bound on $\dim V_x$ bounds $\dim V$ independently of h. We then show that too many choices of x with small $\dim V_x$ would give such a uniform bound.

A bound in terms of the generic rank. Choose bases of V and $\mathcal{S}_t(H)$, and let $M(x)$ be the matrix of the map $T \mapsto T_x$. Its entries are linear in $x_1, \ldots, x_h$. In particular, $\dim V_x$ is the rank of $M(x)$ at the chosen point x. Let R be the rank of this matrix when the coordinates are treated as indeterminates, that is, over $\mathbb{F}_\ell(x_1, \ldots, x_h)$. We claim that

$$\dim V \leq (t+1)!(R + t + 1)^{t+1} \ . \tag{53}$$

The first step is to reduce to a space spanned by individual coefficient forms. For each $\alpha \in \mathbb{N}^h$ with $\sum_i \alpha_i = t + 1$, let E_α have value one on basis tuples with multiplicity vector α, and zero on all other basis tuples. These forms give a basis of $\mathcal{S}_{t+1}(H)$.

Choose positive integer weights $w_1, \ldots, w_h$ so that the numbers $w(\alpha) := \sum_i \alpha_i w_i$ are distinct, and order the coefficient basis by increasing $w(\alpha)$. Put a basis of V in echelon form in this order, with each leading coefficient normalized to one. The leading coefficient forms are then distinct. Introduce an indeterminate z and make the substitution $e_i \mapsto z^{w_i} e_i$. This multiplies the coefficient of E_α by $z^{w(\alpha)}$. Divide each basis form by the power of z corresponding to its leading coefficient. Every remaining coefficient is a polynomial in z, and setting $z = 0$ leaves precisely that leading coefficient form.

The resulting space at $z = 0$ therefore has the same dimension as V and is spanned by distinct E_α. Its generic rank is at most R. Indeed, for indeterminate z, the diagonal substitution and the rescaling of the basis are invertible over $\mathbb{F}_\ell(z)$, so they preserve the generic rank. All $(R+1) \times (R+1)$ minors of the resulting matrix vanish identically in x, z; they still vanish after setting $z = 0$. Thus it suffices to prove (53) for a space spanned by coefficient forms.

To do so, encode each multiplicity vector by the $(t+1)$-element set

$$A_\alpha := \{(i,j) : 1 \le j \le \alpha_i\} \ .$$

We show that more than $(t+1)!(m+t)^{t+1}$ such coefficient forms force a nonzero $m \times m$ minor of $M(x)$. By the distinct-set sunflower bound recalled in the proof of Lemma 6.1, their sets contain a sunflower with $m+t+1$ members. The core is an initial segment in each index i. If an index occurs in the core, at most one petal can contain further copies of that index: two such petals would both contain the next copy, contradicting their disjointness. Discard these at most $t+1$ members and retain m members. Their petal indices are now disjoint from the core indices and from one another.

The column of $M(x)$ belonging to E_α has entry x_i in the row indexed by α with one occurrence of i removed. This follows by evaluating $E_\alpha(x, \cdot, \dots, \cdot)$ on the basis tuple indexing that row. If the sunflower core is nonempty, choose an index i in it and, in each selected column, select this row. Each selected row still contains its own petal, so its entry in every other selected column is zero. The resulting minor is diagonal, with diagonal entries x_i.

If the core is empty, remove one index i_j from the j-th member instead. Since $t+1 \ge 2$, the selected row still contains an index from that member's petal and cannot occur in another column. The minor is again diagonal, now with entries $x_{i_1}, \dots, x_{i_m}$. In both cases its determinant is a nonzero polynomial. Taking $m = R+1$ proves (53).

From ranks at base-field points to generic rank. Suppose that $\dim V_x \le B$ for every $x \in H$. We cannot simply conclude that $R \le B$, since a nonzero minor might vanish at every point of the small base field. We find a point over a field extension and express it using base-field points.

If $R = 0$, every entry of $M(x)$ is identically zero, so $V = 0$. For $R \ge 1$, choose a nonzero $R \times R$ minor. Its degree is R. Over an extension of degree

$$e_0 := \lceil \log_\ell(R+1) \rceil \ ,$$

the field size exceeds R, so Lemma 3.5 gives a point x where this minor is nonzero. Write $x = \sum_{j=1}^{e_0} \theta_j x^{(j)}$, where $\theta_1, \dots, \theta_{e_0}$ is a basis of the extension over $\mathbb{F}_\ell$ and $x^{(j)} \in H$. Linearity of the matrix entries gives

$$M(x) = \sum_{j=1}^{e_0} \theta_j M(x^{(j)}) \ .$$

Each matrix on the right has rank at most B, also over the extension. Subadditivity of rank therefore gives $R \le e_0 B$. Since $e_0 \le 2\sqrt{R}$, we have $R \le 4B^2$. Substituting this into (53) proves the uniform bound

$$\big(\dim V_x \le B \text{ for all } x \in H\big) \quad \Longrightarrow \quad \dim V \le (t+1)!(4B^2+t+1)^{t+1} \ . \tag{54}$$

From many small spaces to a uniform bound. Let

$$S := \{x \in H : \dim V_x < s\} \ .$$

It remains to show that S has density at most $1/\ell + \varepsilon$ when $\dim V$ is large. Suppose otherwise. We will use the density of S to express every $x \in H$ as a sum of a bounded number of elements of S.

The sumset bound in [DeV14] states that, for nonempty $A, B \subseteq H$,

$$|A+B| \ge |A| + |B| - |P| \ , \qquad P := \{z : A + B + z = A + B\} \ .$$

If $A + B \neq H$, its stabilizer P is a proper additive subgroup, so $|P| \leq |H|/\ell$. Apply this with $A = jS$ and $B = S$, where jS denotes the sum of j copies of S. As long as $(j+1)S \neq H$, adding one further copy increases the size by more than $\varepsilon|H|$. This cannot continue for more than a bounded number of steps. In particular, with

$$m_0 := \lceil 1/\varepsilon \rceil + 1 \ ,$$

we have $m_0 S = H$. Here $0 \in S$, so a representation using fewer summands can be padded with zeros.

For $x = x_1 + \cdots + x_{m_0}$ with each $x_j \in S$, multilinearity gives $T_x = \sum_j T_{x_j}$ for every $T \in V$. Consequently,

$$V_x \subseteq V_{x_1} + \cdots + V_{x_{m_0}} \ , \qquad \dim V_x \leq m_0(s-1) \ .$$

This is a uniform bound for all $x \in H$. By (54), we may take

$$K = 1 + (t+1)! \left(4m_0^2(s-1)^2 + t + 1\right)^{t+1} \ .$$

Then $\dim V \geq K$ rules out $|S| > (1/\ell + \varepsilon)|H|$, proving the lemma. □

The dimension estimate now gives a support estimate. The identity $\mu_t(V) = \mathbb{E}_x[\mu_{t-1}(V_x)]$, for $t \geq 2$, lets us average the bounds for forms of one lower order. Most spaces V_x are nonzero, and the preceding lemma ensures that a fraction close to a_ℓ have enough dimension to use the stronger inductive bound.

Lemma 10.10 (Support of high-dimensional spaces)**.** *Fix $t \geq 1$ and $\varepsilon > 0$. There is an integer K, depending only on t, ℓ, ε, such that every $V \leq \mathcal{S}_{t+1}(\mathbb{F}_\ell^h)$ with $\dim V \geq K$ satisfies*

$$\mu_t(V) \geq G_t - \varepsilon \ .$$

For $0 < \varepsilon \leq 1$, the threshold can be chosen to satisfy

$$\log K \leq \exp(O_\ell(t \log(t+1))) \log \frac{2}{\varepsilon} \ . \tag{55}$$

Proof. The probability that $V_x = 0$. The common kernel

$$Z := \{x \in H : V_x = 0\}$$

is a linear subspace of H. Write $w := \operatorname{codim}_H Z$. Every $T \in V$ vanishes when its first argument lies in Z, and symmetry gives the same conclusion for every argument. Thus all forms in V factor through the w-dimensional space H/Z. By Definition 10.6,

$$\begin{aligned} \dim V &\leq \binom{w+t}{t+1} \ , \\ \Pr[V_x = 0] &= \frac{|Z|}{|H|} = \ell^{-w} \ . \end{aligned} \tag{56}$$

For fixed t, the first inequality forces w to grow as $\dim V$ grows. The second then makes the probability of $V_x = 0$ arbitrarily small, uniformly in h.

Induction on t. For $t = 1$, the forms T_x are linear functionals, so $\mu_1(V) = \Pr[V_x \neq 0] = 1 - \ell^{-w}$. Since $G_1 = 1$, this proves the base case.

Suppose $t \geq 2$ and the assertion holds for $t-1$. Choose

$$\eta := \min\left\{\frac{\varepsilon}{4}, \frac{a_\ell}{2}, \frac{\beta_{t-1}}{2(\ell-1)}\right\} \ .$$

Then $\eta < a_\ell$, $3\eta < \varepsilon$ and $G_{t-1} - \eta \geq \beta_{t-1}$. Let s be the dimension threshold in the induction hypothesis with tolerance η. For sufficiently large $\dim V$, Lemma 10.9 and (56) give

$$\begin{aligned} \Pr[\dim V_x \geq s] &\geq a_\ell - \eta \ , \\ \Pr[V_x \neq 0] &\geq 1 - \eta \ . \end{aligned}$$

On every nonzero V_x, choose a nonzero form and apply Lemma 10.8; its support is contained in the support of V_x, so $\mu_{t-1}(V_x) \geq \beta_{t-1}$. When $\dim V_x \geq s$, the induction hypothesis improves this lower bound to $G_{t-1} - \eta$. We therefore count the baseline β_{t-1} on all nonzero spaces, and the additional amount $G_{t-1} - \eta - \beta_{t-1}$ on the large spaces.

Conditioning on the first argument in the definition of μ_t gives

$$\begin{aligned} \mu_t(V) &= \mathbb{E}_x[\mu_{t-1}(V_x)] \\ &\geq (1-\eta)\beta_{t-1} + (a_\ell - \eta)(G_{t-1} - \eta - \beta_{t-1}) \\ &\geq a_\ell G_{t-1} + \ell^{-1}\beta_{t-1} - 3\eta \ . \end{aligned}$$

The parameters in (52) satisfy

$$a_\ell G_{t-1} + \ell^{-1}\beta_{t-1} = \beta_t \left(1 + \frac{t-1}{\ell - 1} + \frac{1}{\ell - 1}\right) = G_t \ .$$

Thus $\mu_t(V) \geq G_t - 3\eta > G_t - \varepsilon$, completing the induction. Both dimension thresholds used above depend only on the fixed parameters, so the resulting K is independent of h.

To track its size, write $K_{t,\ell}(\varepsilon)$ for a valid threshold. In the base case we can take

$$K_{1,\ell}(\varepsilon) = 1 + \binom{\lceil \log_\ell(1/\varepsilon) \rceil + 1}{2} \ .$$

For $t \geq 2$, put $w_0 := \lceil \log_\ell(1/\eta) \rceil$ and $s := K_{t-1,\ell}(\eta)$. The explicit threshold in Lemma 10.9 and (56) allow

$$K_{t,\ell}(\varepsilon) \leq 2 + \binom{w_0 + t}{t+1} + (t+1)! \left(4(\lceil 1/\eta \rceil + 1)^2 s^2 + t + 1\right)^{t+1} \ .$$

Since $\log(1/\eta) \leq \log(1/\varepsilon) + O_\ell(t)$, taking logarithms and iterating this recurrence gives (55). Each step multiplies the preceding logarithmic bound by $O(t)$; their product is $\exp(O(t \log(t+1)))$. □

The starting codes. We now choose a particular space of forms. The codewords in the list will be differences of pure powers, and their differences from the center of the ball will be pure powers. This will give a list whose errors have block weight β_t and whose pairwise codeword differences have block weight $\delta_t > \beta_t$. We will also ensure that the space contains no nonzero pure power, so Lemma 10.8 gives the stronger minimum-weight bound throughout the code.

For $u \in H$, identify u with the functional $z \mapsto u^{\mathsf{T}} z$, and put

$$\begin{aligned} U_h &:= \{(1, u_2, \ldots, u_h) : u_i \in \{0,1\}\} \ , \\ u_0 &:= (1, 0, \ldots, 0) \ , \\ T_u &:= u^{\otimes(t+1)} \ , \\ B_h &:= \operatorname{span}_{\mathbb{F}_\ell}\{T_u - T_{u_0} : u \in U_h\} \leq \mathcal{S}_{t+1}(H) \ . \end{aligned} \tag{57}$$

The restriction $u_i \in \{0,1\}$ is used over every prime field. We identify B_h with its image under $T \mapsto \Phi_T$. This is a linear code with one $H^\vee$-valued block for each point of H^t, and distances are measured by the fraction of blocks on which two words differ. Its blocklength is $|H|^t = \ell^{ht}$. The following lemma records the dimension, the large list, and the support bound that we will preserve when passing to a much shorter scalar-valued code.

Lemma 10.11 (Tensor seed)**.** *For $h > t+1$, the space B_h has dimension*

$$M_h := \dim B_h = \sum_{j=1}^{t+1} \binom{h-1}{j} = \Theta(h^{t+1}) \ . \tag{58}$$

The evaluated code has relative minimum block distance exactly δ_t, and the ball of relative block radius β_t centered at $\Phi_{T_{u_0}}$ contains at least 2^{h-1} distinct codewords. For every $\eta > 0$ there is a fixed $k \geq 2$, independent of h, such that every k-dimensional $V \leq B_h$ satisfies $\mu_t(V) \geq G_t - \eta$ whenever $h \geq \max\{t+2, k+1\}$. For $0 < \eta \leq 1$, the integer k can satisfy the bound in (55) *with tolerance η.*

Proof. The dimension. Consider the coefficients of T_u as functions of $u \in U_h$. Each coefficient is a product of $t+1$ coordinates of u. Since $u_1 = 1$ and $u_i^j = u_i$ for $u_i \in \{0,1\}$ and $j \geq 1$, these products are precisely the squarefree monomials in $u_2, \ldots, u_h$ of degree at most $t+1$. Every such monomial occurs: use its indices once, and fill the remaining slots with the first index.

These monomials are linearly independent as functions on $\{0,1\}^{h-1}$ over every field. Indeed, evaluating a linear relation at the indicator of a subset A gives the sum of the coefficients indexed by subsets of A. Starting with $A = \emptyset$ and proceeding by increasing $|A|$ forces each coefficient to be zero. The matrix whose columns are the coefficient vectors of T_u therefore has rank $\sum_{j=0}^{t+1} \binom{h-1}{j}$.

To pass from the span of the T_u to B_h, observe that $T_u(e_1, \ldots, e_1) = 1$ for every u. The differences defining B_h lie in the kernel of this coefficient. Conversely, if $\sum_u c_u T_u$ is in the kernel, then $\sum_u c_u = 0$, and

$$\sum_u c_u T_u = \sum_u c_u (T_u - T_{u_0}) \in B_h \ .$$

Thus B_h is exactly this kernel, of codimension one in the span of the T_u. This proves (58).

Excluding pure powers. For every $T \in B_h$,

$$\begin{aligned} &T(e_1, \ldots, e_1) = 0 \ , \\ &T(e_j, \ldots, e_j) = T(e_j, e_1, \ldots, e_1) \qquad (1 \leq j \leq h) \ . \end{aligned}$$

The second identity holds for each T_u because $u_j^{t+1} = u_j$, and hence also for their differences and linear combinations. If a nonzero T were $c\, v^{\otimes(t+1)}$, the first identity would give $v(e_1) = 0$. The second would then give $c\, v(e_j)^{t+1} = 0$ for every j, so $v = 0$, a contradiction. By Lemma 10.8, every nonzero codeword therefore has relative block weight at least δ_t.

The list and the exact distance. Consider the word $\Phi_{T_{u_0}}$. For each $u \in U_h$, the code contains $\Phi_{T_{u_0} - T_u}$, whose error relative to that word is Φ_{T_u}. Under the identification of $H^\vee$ with coordinate vectors,

$$\Phi_{T_u}(x_1, \ldots, x_t) = u \prod_{j=1}^{t} (u^{\mathsf{T}} x_j) \ .$$

Each factor is nonzero with probability a_ℓ, independently across the arguments. The error thus has relative block weight $a_\ell^t = \beta_t$.

For distinct $u, v \in U_h$, their first coordinates are one, so they cannot be scalar multiples. The pair $(u^{\mathsf{T}} x_j, v^{\mathsf{T}} x_j)$ is therefore uniform on $\mathbb{F}_\ell^2$ for each uniform x_j. Consequently, the supports of Φ_{T_u} and Φ_{T_v} have relative sizes β_t, β_t, and their intersection has relative size β_t^2. On the intersection the two values are nonzero multiples of linearly independent vectors and cannot cancel. Hence

$$\mu_t(T_u - T_v) = 2\beta_t - \beta_t^2 = \delta_t \ .$$

This proves that all $|U_h| = 2^{h-1}$ listed codewords are distinct, and also supplies a word attaining the minimum-distance lower bound.

The generalized support bound. Fix $\eta > 0$ and choose k at least two and at least the dimension threshold in Lemma 10.10. Since $M_h \geq h - 1$, the space B_h has dimension at least k whenever $h \geq k + 1$. Every k-dimensional subspace of B_h then has relative block support at least $G_t - \eta$, as required. □

Thus the codewords in the list are at relative block distance β_t from $\Phi_{T_{u_0}}$, the minimum relative block distance is δ_t, and every k-dimensional subspace has relative block support at least $G_t - \eta$. The last quantity will give the generalized weight needed for MCA. The next step converts these three block-support estimates into ordinary Hamming-weight estimates for scalar-valued codes.

10.3 Proof of Theorem 10.1

The starting code has 2^{h-1} nearby codewords and dimension $M_h = \Theta(h^{t+1})$, but its blocklength is ℓ^{ht} and its symbols lie in $H^\vee$. We will construct a scalar-valued code by applying $N_h = \Theta(M_h)$ linear functionals to the forms in B_h. The new blocklength is $\Theta(h^{t+1})$ and the dimension is unchanged. Since t is fixed, the list size 2^{h-1} is superpolynomial in the new blocklength, and the rate is bounded away from zero.

The functionals are chosen so that a form of relative block weight s gives a nonzero scalar coordinate with probability

$$F(s) = b(1 - e^{-\lambda s}) \ , \qquad b := 1 - 1/q \ .$$

The same function gives a lower bound on the probability that a coordinate belongs to the support of a subspace. Thus the three support estimates from Lemma 10.11 will become a minimum distance close to $F(\delta_t)$, a list radius at most $F(\beta_t)$, and a k-th generalized weight at least $F(G_t - \eta)$, up to sampling errors.

To see why these quantities can meet the desired conditions, first consider the limiting choice $b = 1$ and $\lambda = -\log(1 - d)/\delta_t$. The three corresponding values are

$$\begin{aligned} F(\delta_t) &= d \ , \\ F(\beta_t) &= 1 - (1 - d)^{\beta_t/\delta_t} \ , \\ F(G_t) &= 1 - (1 - d)^{G_t/\delta_t} \ . \end{aligned}$$

The support ratios satisfy

$$\frac{\beta_t}{\delta_t} = \frac{1}{2 - \beta_t} \longrightarrow \frac{1}{2} \ , \qquad \frac{G_t}{\delta_t} = \frac{1 + t/(\ell - 1)}{2 - \beta_t} \longrightarrow \infty \qquad (t \to \infty) \ . \tag{59}$$

The list radius therefore approaches $J_2(d)$, while the generalized-weight lower bound approaches one. Since $p > J_2(d)$ and $\Psi(\gamma) < 1$, a sufficiently large fixed t leaves a positive margin in both desired inequalities. The proof below chooses a large fixed q and small sampling tolerances within these margins. It then constructs the scalar code and appends zero coordinates to make its relative distance tend to exactly d.

Proof. Choosing the parameters. Fix d, p, γ as in (42), and let ξ be as in (43). Choose an integer $t \geq 1$ such that

$$\beta_t \leq \frac{\xi}{8\log(2/\xi)} \ ,$$
$$t \geq \frac{2(\ell-1)}{d}\log\frac{8}{\xi} \ .$$

Since $\beta_t = (1-1/\ell)^t$, we can take

$$t = O_\ell\left(\xi^{-1}\log\frac{2}{\xi}\right) \ . \tag{60}$$

The first condition gives $\beta_t/\delta_t \leq 1/2 + \beta_t/2$, and hence

$$1-(1-d)^{\beta_t/\delta_t} \leq J_2(d) + \frac{\beta_t}{2}\log\frac{1}{1-d} \leq J_2(d) + \frac{\xi}{16} \ .$$

The second condition and (59) give $G_t/\delta_t \geq t/(2(\ell-1))$, so

$$1-(1-d)^{G_t/\delta_t} \geq 1 - \frac{\xi}{8} \ .$$

Thus both desired inequalities have margins comparable to ξ.

Set

$$\varepsilon := \frac{\xi^2}{4096}, \qquad \eta := \frac{\varepsilon\delta_t}{\log(2/\xi)}, \qquad d' := d + 2\varepsilon \ ,$$

and let q be the least power of ℓ at least $1/\varepsilon$. Then $q = O_\ell(\xi^{-2})$. Put $b := 1 - 1/q$ and define

$$\lambda := -\frac{\log(1-d'/b)}{\delta_t} \ ,$$
$$F(s) := b(1-e^{-\lambda s}) = b\left(1-(1-d'/b)^{s/\delta_t}\right) \qquad (s \geq 0) \ . \tag{61}$$

We have $b \geq 1-\varepsilon$ and $1-d'/b \geq \xi/2$, so λ is well-defined and $F(\delta_t) = d'$. The sampled code will have distance slightly above d; appending zeros will then adjust its relative distance without identifying codewords.

We verify quantitatively that the margins survive. For $r > 0$, write $H_r(b,x) := b(1-(1-x/b)^r)$. On the rectangle $1-\varepsilon \leq b \leq 1$, $d \leq x \leq d'$, we have $x \geq \xi$ and $1 - x/b \geq \xi/2$. Its derivatives satisfy $|\partial_x H_r| \leq 2/\xi$ and $|\partial_b H_r| \leq 3/\xi$, uniformly in r. Indeed, put $z := 1 - x/b \in [\xi/2, 1-\xi]$. Since $\xi \leq 1/2$, concavity gives $z\log(1/z) \geq \xi/2$, and hence

$$0 \leq \partial_x H_r = r z^{r-1} \leq \frac{1}{z\log(1/z)} \leq \frac{2}{\xi} \ .$$

Also, $\partial_b H_r = 1 - z^r - (1-z)rz^{r-1}$, so $|\partial_b H_r| \leq 1 + 2/\xi \leq 3/\xi$. Thus

$$|H_r(b,d') - H_r(1,d)| \leq \frac{7\varepsilon}{\xi} \ .$$

Moreover, $F'(s) \leq \lambda \leq \log(2/\xi)/\delta_t$, so replacing G_t by $G_t - \eta$ loses at most ε. Consequently,

$$d' - \varepsilon > d \ ,$$

$$F(\beta_t) + \varepsilon \le J_2(d) + \frac{\xi}{16} + \frac{7\varepsilon}{\xi} + \varepsilon < p - \frac{\xi}{2} \ ,$$
$$\frac{d}{d' + \varepsilon}\big(F(G_t - \eta) - \varepsilon\big) \ge 1 - \frac{\xi}{8} - \frac{10\varepsilon}{\xi} - 2\varepsilon > \Psi(\gamma) + \frac{\xi}{2} \ . \tag{62}$$

For the last line we used $d/(d + 3\varepsilon) \ge 1 - 3\varepsilon/\xi$ and $\Psi(\gamma) \le 1 - \xi$. It includes the decrease in relative generalized weight caused by padding.

Take $k := \max\{2, K_{t,\ell}(\eta)\}$, with the threshold from Lemma 10.10. Since

$$\log \frac{1}{\eta} = O_\ell\left(t + \log \frac{2}{\xi}\right),$$

(55) and (60) give the bound on k in (46). All these parameters are fixed while h grows. We may start at

$$h_0 := \lceil \max\{k + 2,\ 2(t + 2),\ 4/\xi + 2,\ q\} \rceil \ . \tag{63}$$

Passing to the fixed field $\mathbb{F}_q$. Put $B_h^{(q)} := B_h \otimes_{\mathbb{F}_\ell} \mathbb{F}_q$ and extend the forms multilinearly to $\mathbb{F}_q^h$. We continue to evaluate the first t arguments at points of $H = \mathbb{F}_\ell^h$. Each evaluation block is now a linear functional on $\mathbb{F}_q^h$. We use Φ_T and μ_t for this evaluation and its relative block support.

The block-support bounds survive this extension. Indeed, after expansion in an $\mathbb{F}_\ell$-basis of $\mathbb{F}_q$, an evaluation block is zero exactly when all its components are zero. The shortening argument in Lemma 10.4 therefore applies with whole blocks as the coordinate sets, and preserves the block generalized weights and dimension. For every $h \ge h_0$, we have $M_h \ge h - 1 \ge k$, and consequently

$$\dim B_h^{(q)} = M_h \ ,$$
$$\min_{0 \ne T \in B_h^{(q)}} \mu_t(T) = \delta_t \ ,$$
$$\mu_t(V) \ge G_t - \eta \qquad \text{for every } V \le B_h^{(q)} \text{ with } \dim V = k \ .$$

The distinguished forms T_u and their differences remain defined over $\mathbb{F}_\ell$, so the list from Lemma 10.11 is unchanged. This first scalar extension allows $b = 1 - 1/q$ to be close to one; the further extensions used to reduce the MCA error will come at the end.

Constructing a scalar coordinate. Draw J from the Poisson distribution of mean λ. Conditional on J, choose independently uniform positions $X_j = (x_{j,1}, \dots, x_{j,t}) \in H^t$ and uniform vectors $y_j \in \mathbb{F}_q^h$, for $1 \le j \le J$. Define

$$L(T) := \sum_{j=1}^{J} T(x_{j,1}, \dots, x_{j,t}, y_j) \ .$$

This is a $\mathbb{F}_q$-linear functional on the full space of scalar-extended symmetric forms. Each summand first selects an evaluation block $\Phi_T(X_j)$ and then evaluates that functional at y_j. The output $L(T)$ is a single scalar, even though the starting blocks are vector-valued.

Let T have relative block weight s. Conditional on the positions, if every sampled block is zero, then $L(T) = 0$. If at least one sampled block is nonzero, its evaluation at the independent uniform y_j is uniform on $\mathbb{F}_q$. The sum is then uniform on $\mathbb{F}_q$ as well, and is nonzero with probability b. The probability that all positions miss the support is

$$\mathbb{E}[(1 - s)^J] = \sum_{j=0}^{\infty} e^{-\lambda} \frac{\lambda^j}{j!}(1 - s)^j = e^{-\lambda s} \ .$$

This explains both the Poisson sampling and the function F: they give the exact identity

$$\Pr[L(T) \neq 0] = b(1 - e^{-\lambda s}) = F(s) \ .$$

For a subspace V with relative block support s, a scalar coordinate belongs to the support of its image precisely when $L|_V \neq 0$. Conditional on the sampled positions, this restriction is the image of the uniform tuple $(y_1, \ldots, y_J)$ under a linear map into $V^\vee$. It is therefore uniform on the image of that map. If every position misses the support of V, the image is zero. If at least one position hits the support, the image has positive dimension, and a uniform element of it is nonzero with probability at least $1 - 1/q = b$. Averaging over the positions gives

$$\Pr[L|_V \neq 0] \geq F(s) \ .$$

Thus a lower bound on block support gives a lower bound on the expected ordinary support of the scalar image, for both words and subspaces.

Choosing all coordinates simultaneously. Take $N_h := \lceil cM_h \rceil$ independent copies $L_1, \ldots, L_{N_h}$, where c is a fixed constant to be chosen, and put

$$A_h(T) := (L_1(T), \ldots, L_{N_h}(T)) \ .$$

The scalar code is $A_h(B_h^{(q)}) \subseteq \mathbb{F}_q^{N_h}$. It is spanned by the words $A_h(T_u - T_{u_0})$. Define

$$\widehat{\mathcal{L}}_h := \{A_h(T_{u_0} - T_u) : u \in U_h\}, \qquad \widehat{z}_h := A_h(T_{u_0}) \ .$$

These codewords have differences $\widehat{z}_h - A_h(T_{u_0} - T_u) = A_h(T_u)$ from $\widehat{z}_h$. The map A_h is defined on the full space of forms, so $\widehat{z}_h$ and these differences are defined even though T_u need not belong to $B_h^{(q)}$.

We must preserve the distance, the generalized weight, and this list with one choice of the rows. We require

$$\begin{aligned}
\left| \frac{\mathsf{wt}(A_h(T))}{N_h} - F(\mu_t(T)) \right| &\leq \varepsilon && \text{for every } T \in B_h^{(q)} \ , \\
\frac{|\mathsf{supp}(A_h(V))|}{N_h} &\geq F(\mu_t(V)) - \varepsilon && \text{for every } V \leq B_h^{(q)},\ \dim V = k \ , \\
\frac{\mathsf{wt}(A_h(T_u))}{N_h} &\leq F(\beta_t) + \varepsilon && \text{for every } u \in U_h \ .
\end{aligned}$$

For a fixed form, its output weight is the sum of independent Bernoulli variables of mean $F(\mu_t(T))$. For a fixed subspace, its output support is the sum of independent Bernoulli variables of mean at least $F(\mu_t(V))$. Hoeffding's inequality bounds the failure probability by $2e^{-2N_h\varepsilon^2}$ for each two-sided estimate and by $e^{-2N_h\varepsilon^2}$ for each one-sided estimate.

There are q^{M_h} forms in $B_h^{(q)}$, at most q^{kM_h} subspaces of dimension k, and $|U_h| = 2^{h-1} \leq q^{M_h}$ distinguished error forms. A union bound gives a total failure probability of at most

$$\big(2q^{M_h} + q^{kM_h} + q^{M_h}\big)e^{-2N_h\varepsilon^2} \leq 4\exp\big((k\log q - 2c\varepsilon^2)M_h\big) \ .$$

Take

$$c := \frac{k\log q + \log 16}{2\varepsilon^2} = O\left(\frac{k\log q + 1}{\xi^4}\right) \ .$$

The failure probability is at most $4 \cdot 16^{-M_h} < 1$ for every $h \geq h_0$. For each such h, fix rows satisfying all three estimates and define the resulting deterministic code

$$\widehat{B}_h := A_h(B_h^{(q)}) \subseteq \mathbb{F}_q^{N_h} \ .$$

Translating the support bounds. Every nonzero $T \in B_h^{(q)}$ has $\mu_t(T) \geq \delta_t$. Since F is increasing and $F(\delta_t) = d'$, the first estimate gives $\mathsf{wt}(A_h(T))/N_h \geq d' - \varepsilon > 0$. Hence A_h is injective on $B_h^{(q)}$, preserving its dimension and the distinctness of the 2^{h-1} words in $\widehat{\mathcal{L}}_h$. On the other hand, for distinct $u, v \in U_h$, the form $T_u - T_v$ has block weight exactly δ_t, so its image has relative weight at most $d' + \varepsilon$. These two observations bound the minimum distance on both sides.

Injectivity also ensures that every k-dimensional subspace of $\widehat{B}_h$ is the image of a k-dimensional subspace of $B_h^{(q)}$. Its block support is at least $G_t - \eta$, so the second estimate bounds the ordinary support of its image below by $N_h(F(G_t - \eta) - \varepsilon)$. Minimizing over these subspaces gives the ordinary k-th generalized Hamming weight. Finally, the third estimate places all words in $\widehat{\mathcal{L}}_h$ within relative distance $F(\beta_t) + \varepsilon$ of $A_h(T_{u_0})$.

Writing $\delta_h := \delta(\widehat{B}_h)$, we have therefore obtained

$$\begin{aligned}
&\dim \widehat{B}_h = M_h \ , \\
&d' - \varepsilon \leq \delta_h \leq d' + \varepsilon \ , \\
&\frac{d_k\Big(\widehat{B}_h\Big)}{N_h} \geq F(G_t - \eta) - \varepsilon \ , \\
&L_{\widehat{B}_h}(F(\beta_t) + \varepsilon) \geq 2^{h-1} \ .
\end{aligned} \tag{64}$$

All weights and radii in this display are ordinary Hamming weights and radii of a scalar-valued code.

Adjusting the distance by padding. Put $D_h := \Delta\Big(\widehat{B}_h\Big) = \delta_h N_h$ and

$$n_h := \lceil D_h / d \rceil \ .$$

Since $\delta_h \geq d' - \varepsilon > d$, we have $n_h > N_h$. Define $\mathcal{C}_{n_h}$ by appending $n_h - N_h$ zero coordinates to every word of $\widehat{B}_h$. Denote the corresponding padded list and word by $\mathcal{L}_h$ and z_h, respectively.

Padding leaves the dimension, the absolute minimum distance, every absolute generalized weight, and the distances from the words in $\widehat{\mathcal{L}}_h$ to $\widehat{z}_h$ unchanged. Each relative weight or radius is consequently multiplied by N_h/n_h. Since $n_h = D_h/d + O(1)$,

$$\begin{aligned}
\delta(\mathcal{C}_{n_h}) = \frac{D_h}{n_h} &\longrightarrow d \ , \\
\frac{N_h}{n_h} = \frac{d}{\delta_h} + O(N_h^{-1}) &\geq \frac{d}{d' + \varepsilon} - O(N_h^{-1}) \ .
\end{aligned}$$

In particular,

$$\max_{\mathbf{c} \in \mathcal{L}_h} \frac{\Delta(z_h, \mathbf{c})}{n_h} \leq \frac{N_h}{n_h}\big(F(\beta_t) + \varepsilon\big) \leq F(\beta_t) + \varepsilon < p \ .$$

The generalized weight satisfies the parallel identity

$$\frac{d_k(\mathcal{C}_{n_h})}{n_h} = \frac{N_h}{n_h} \cdot \frac{d_k\Big(\widehat{B}_h\Big)}{N_h} \ .$$

More precisely, $n_h \le D_h/d + 1$ and $\delta_h \ge d$ give

$$\frac{N_h}{n_h} \ge \frac{d}{\delta_h + d/N_h} \ge \frac{d}{d' + \varepsilon} - \frac{1}{N_h} \ .$$

Combining this with (64) and (62), and using $N_h \ge M_h \ge h - 1 \ge 4/\xi$, yields

$$\frac{d_k(\mathcal{C}_{n_h})}{n_h} > \Psi(\gamma) + \frac{\xi}{2} - \frac{1}{N_h} \ge \Psi(\gamma) + \frac{\xi}{4} \ .$$

Thus padding achieves the prescribed distance while keeping $\mathcal{L}_h$ within relative distance p of z_h and preserving the generalized-weight margin for every $h \ge h_0$.

The rate and list size. The sequence δ_h lies in the fixed interval $[d' - \varepsilon, d' + \varepsilon]$. Pass to a subsequence along which $\delta_h \to \delta_* > d$. Since $N_h/M_h \to c$, along this subsequence

$$\frac{\dim \mathcal{C}_{n_h}}{n_h} = \frac{M_h}{N_h} \cdot \frac{N_h}{n_h} \longrightarrow R := \frac{d}{c\delta_*} > 0 \ .$$

The sampled and padded lengths differ by only a bounded factor, so

$$n_h = \Theta(N_h) = \Theta(M_h) = \Theta(h^{t+1}) \ .$$

These lengths are unbounded, and $|\mathcal{L}_h| = 2^{h-1}$ gives

$$L_{\mathcal{C}_{n_h}}(p) \ge 2^{h-1} \ge 2^{c_1 n_h^{1/(t+1)}}$$

for a fixed $c_1 > 0$, uniformly for $h \ge h_0$. This is superpolynomial in the final length. To obtain the stated form of (45), take $\beta := 1/(2(t+1))$ and $c_0 := c_1 \beta \log 2$. The inequality $\log n \le n^\beta/\beta$ then gives the required bound for every length in the constructed sequence. Equation (60) gives the claimed bound on β^{-1}.

Scalar extension and MCA. Set $\eta_0 := \xi/4$. For every $h \ge h_0$, the generalized-weight estimate above gives

$$\frac{d_k(\mathcal{C}_{n_h})}{n_h} \ge \Psi(\gamma) + \eta_0 \ .$$

Consequently,

$$1 - \frac{n_h \, \Psi(\gamma)}{d_k(\mathcal{C}_{n_h})} \ge \frac{\eta_0}{\Psi(\gamma) + \eta_0} > 0 \ . \tag{65}$$

This is the hypothesis of Lemma 10.5, with a denominator bounded away from zero independently of h.

For any finite extension $\mathbb{K}/\mathbb{F}_q$, define

$$\mathcal{C}_{n_h,\mathbb{K}} := \mathcal{C}_{n_h} \otimes_{\mathbb{F}_q} \mathbb{K} \ .$$

By Lemma 10.4, the dimension, minimum distance and generalized weights are unchanged. The words in $\mathcal{L}_h$ and the word z_h lie over $\mathbb{F}_q$, so their distances are unchanged over $\mathbb{K}$. Applying Lemma 10.5 with (65) yields

$$\epsilon_{\mathsf{MCA}}^{\mathcal{C}_{n_h,\mathbb{K}}}(\gamma) \le \frac{\mathsf{T}_k(n_h, q) + 3}{\left(1 - \frac{n_h \Psi(\gamma)}{d_k(\mathcal{C}_{n_h})}\right) |\mathbb{K}|} \le \frac{C_0 n_h^k}{|\mathbb{K}|} \ .$$

Since $\Psi(\gamma) + \eta_0 \le 1$ and $\eta_0 = \xi/4$, the lower bound in (65) is at least $\xi/4$. Also $\mathsf{T}_k(n, q) \le kq^{k-1}n^k$, so we may take

$$C_0 := \frac{4}{\xi}\big(kq^{k-1} + 3\big) \ .$$

This proves the MCA bound together with the rate, distance and list-size claims.

The size requirements. For $h \ge 2(t+2)$, the dimension formula gives $M_h \le (t+1)h^{t+1}$. Since $N_h \le (c+1)M_h$ and $n_h \le N_h/d + 1$, we have

$$n_h \le Ah^{t+1}, \qquad A := \frac{2(c+1)(t+1)}{\xi} \ .$$

Thus a valid length threshold is

$$n_0 := \lceil Ah_0^{t+1} \rceil \ .$$

The bounds on t, k, q, c, h_0 give

$$\log\log n_0 = O_\ell(t\log(t+1)) = O_\ell\left(\xi^{-1}\log^2\frac{2}{\xi}\right),$$

which proves (46). We retain only lengths at least n_0 before taking the rate-convergent subsequence. All construction estimates then hold at every length in this sequence.

Finally, $n \ge n_0 \ge \max\{2, k, q, 4/\xi\}$ implies $C_0 \le n^{k+3}$. Taking $e = \lceil (2k+5)\log_q n \rceil$ therefore gives

$$n^{2k+5} \le |\mathbb{K}| < qn^{2k+5} \le n^{2k+6}, \qquad \epsilon_{\mathsf{MCA}}^{\mathcal{C}_{n,\mathbb{K}}}(\gamma) \le \frac{n^{2k+3}}{|\mathbb{K}|} \le n^{-2} \ .$$

All parameter choices depend only on ℓ, d, p, γ, so the constants are independent of the final length and extension field size. □

10.4 Proof of Corollary 10.2

Proof. Apply Theorem 10.1 with the fixed parameters ℓ, d, p, γ, and retain its notation $q, k, n_0, c_0, C_0, \beta$. For each length n in that theorem, choose an extension $\mathbb{K}_n/\mathbb{F}_q$ of degree

$$e_n := \lceil (2k+5)\log_q n \rceil, \qquad |\mathbb{K}_n| = q^{e_n} \ ,$$

and put $\mathcal{D}_n := \mathcal{C}_n \otimes_{\mathbb{F}_q} \mathbb{K}_n$. These fields have characteristic ℓ, and

$$n^{2k+5} \le |\mathbb{K}_n| < qn^{2k+5} \le n^{2k+6}$$

for every length $n \ge n_0$ in the sequence. Thus the field-size bounds hold with $A := 2k + 6$.

The rate and relative distance tend to $R > 0$ and d by Theorem 10.1. Moreover, (45) gives

$$\epsilon_{\mathsf{MCA}}^{\mathcal{D}_n}(\gamma) \le \frac{C_0 n^k}{|\mathbb{K}_n|} \le n^{-2} \le n^{-1}$$

for every length $n \ge n_0$ in the sequence.

It remains to compare the list size with an arbitrary polynomial Q. The normalized MCA error satisfies

$$0 \le |\mathbb{K}_n|\epsilon_{\mathsf{MCA}}^{\mathcal{D}_n}(\gamma) \le C_0 n^k \ ,$$

and s^{-1} is fixed. Bounding each monomial of Q in absolute value therefore gives constants $B_Q > 0$ and $D_Q \geq 0$ such that

$$\left|Q\big(n, |\mathbb{K}_n|\epsilon_{\mathsf{MCA}}^{\mathcal{D}_n}(\gamma), s^{-1}\big)\right| \leq B_Q n^{D_Q} \quad .$$

On the other hand, the list-size bound in (45) reads

$$L_{\mathcal{D}_n}(p) \geq 2^{c_0 n^\beta \log_2 n} = n^{c_0 n^\beta} \quad .$$

Since $c_0, \beta > 0$, this exceeds $B_Q n^{D_Q}$ for all sufficiently large n, proving (47). The family was chosen before Q, and all its parameters depend only on ℓ, d, p, γ. □

A Numerical estimate for Lemma 4.11

Lemma A.1. *In the setting of Lemma 4.11, if* $\tau = \frac{1-\gamma+\sqrt{1-\delta_{\mathcal{C}}}}{2}$*, then*

$$\frac{\tau - (1-\delta_{\mathcal{C}})}{\tau^2 - (1-\delta_{\mathcal{C}})} < 5 \ .$$

Consequently, the integer number of curves in Lemma 4.11 is at most 4*.*

Proof. For this numerical calculation, put $a := 1-\gamma$ and $u := \sqrt{1-\delta_{\mathcal{C}}}$. Now let $\tau = \frac{a+u}{2} \in (u,1)$. Then $\tau^2 - u^2 = \frac{(a+3u)(a-u)}{4}$ and $\tau - u^2 = \frac{a+u-2u^2}{2}$, so

$$\Phi(a,u) \ := \ \frac{\tau - u^2}{\tau^2 - u^2} \ = \ \frac{2\left(a+u-2u^2\right)}{(a+3u)(a-u)} \ ,$$

which is decreasing in a (because τ increases with a). By (3) it therefore suffices to bound Φ at $a = a_0 := \max\{u^{2/3}, \frac{1+u^2}{2}\}$; the two expressions in the maximum coincide exactly when $a_0^3 - 2a_0 + 1 = 0$ with $a_0 \neq 1$, i.e. at $a_0 = \frac{\sqrt{5}-1}{2}$ and $u = u^* := \left(\frac{\sqrt{5}-1}{2}\right)^{3/2}$, with $a_0 = \frac{1+u^2}{2}$ for $u \leq u^*$ and $a_0 = u^{2/3}$ for $u \geq u^*$.

Case $u \leq u^*$. Using $a_0 + u - 2u^2 = \frac{(1+3u)(1-u)}{2}$, $a_0 + 3u = \frac{1+6u+u^2}{2}$ and $a_0 - u = \frac{(1-u)^2}{2}$,

$$\Phi(a_0,u) = \frac{4(1+3u)}{(1+6u+u^2)(1-u)} = \frac{4(1+3u)}{1+5u-5u^2-u^3} \ .$$

Hence $\Phi(a_0,u) < 5$ is equivalent to $f(u) := 1 + 13u - 25u^2 - 5u^3 > 0$. We have $(u^*)^2 = \sqrt{5} - 2 < 1/4$, so $u^* < 1/2$. Moreover, $f''(u) = -50 - 30u < 0$ on $[0, 1/2]$, so f is concave there and

$$f(u) \geq \min\{f(0), f(1/2)\} = \frac{5}{8} > 0 \qquad (0 \leq u \leq u^*) \ .$$

Case $u \geq u^*$. Substituting $u = \sigma^3$, so $a_0 = \sigma^2$ and $\sigma \geq \sigma^* := (u^*)^{1/3} = \sqrt{\frac{\sqrt{5}-1}{2}}$,

$$\Phi(a_0,u) = \frac{2\left(\sigma^2 + \sigma^3 - 2\sigma^6\right)}{(\sigma^2+3\sigma^3)(\sigma^2-\sigma^3)} = \frac{2\left(1+\sigma-2\sigma^4\right)}{\sigma^2+2\sigma^3-3\sigma^4} \ ,$$

and $\Phi(a_0,u) < 5$ is equivalent to $-11\sigma^4 + 10\sigma^3 + 5\sigma^2 - 2\sigma - 2 > 0$, that is, after factoring out the root $\sigma = 1$, to $(1-\sigma)\left(11\sigma^3 + \sigma^2 - 4\sigma - 2\right) > 0$. The cubic $g(\sigma) := 11\sigma^3 + \sigma^2 - 4\sigma - 2$ is increasing for $\sigma \geq 1/3$. Since $(\sigma^*)^2 = \frac{\sqrt{5}-1}{2} > 9/16$, we have $\sigma^* > 3/4$, and hence

$$g(\sigma) \geq g(3/4) = \frac{13}{64} > 0 \qquad (\sigma^* \leq \sigma < 1) \ .$$

In both cases $\Phi(a,u) < 5$, so $L \leq 4$. □

B Numerical estimate for Theorem 7.1

We first isolate the only numerical ingredient of the proof.

Lemma B.1 (Numerical estimate)**.** *Let* $\tau \in (1-\delta_{\mathcal{C}}, \frac{1-\gamma+\sqrt{1-\delta_{\mathcal{C}}}}{2}]$ *and put* $X := \dfrac{\tau-(1-\delta_{\mathcal{C}})}{\tau(1-\gamma-\tau)}$*. Then, under* (3)*,*

$$\frac{J_h(\delta_{\mathcal{C}})}{\delta_{\mathcal{C}}}\left(\frac{11}{5}+X\right) \le \frac{31}{5} \ .$$

Proof. In this numerical calculation only, write $a := 1-\gamma$, $u := (1-\delta_{\mathcal{C}})^{1/2}$ and $b := (1-\delta_{\mathcal{C}})^{1/h}$, and set

$$\Gamma \ := \ 1+b+b^2+\cdots+b^{h-1} \ = \ \frac{1-b^h}{1-b} \ = \ \frac{\delta_{\mathcal{C}}}{1-b} \ \ge \ 1 \ . \tag{66}$$

Reduction to $\tau = \frac{a+u}{2}$. On (u^2, a) we have

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\log\frac{\tau-u^2}{\tau(a-\tau)} = \frac{1}{\tau-u^2}-\frac{1}{\tau}+\frac{1}{a-\tau} = \frac{u^2}{\tau(\tau-u^2)}+\frac{1}{a-\tau} \ > \ 0 \ ,$$

so X increases with τ and it suffices to treat $\tau = \frac{a+u}{2}$, where $\tau-u^2 = \frac{a+u-2u^2}{2}$ and $a-\tau = \frac{a-u}{2}$, so that

$$X \ \le \ \frac{2\left(a+u-2u^2\right)}{a^2-u^2} \ .$$

Reduction to the extreme value of a*.* Writing $\chi := u-2u^2$,

$$\frac{\mathrm{d}}{\mathrm{d}a}\frac{2(a+\chi)}{a^2-u^2} = \frac{-2\left(a^2+2\chi a+u^2\right)}{\left(a^2-u^2\right)^2} \ \le \ 0 \ ,$$

since $a^2+2\chi a+u^2$ has discriminant $4(\chi^2-u^2) = 4(\chi-u)(\chi+u) = -16u^3(1-u) \le 0$. Hence, by (3), X is bounded by the same expression evaluated at $a = a_0 := \max\{u^{2/3}, \frac{1+u^2}{2}\}$.

Reduction to a one-variable inequality. Put $\sigma := u^{1/3}$, so $u = \sigma^3$ and $1-\delta_{\mathcal{C}} = u^2 = \sigma^6$. Since $h \ge 3$ we have $b \ge (1-\delta_{\mathcal{C}})^{1/3} = \sigma^2$, hence by (66)

$$\Gamma \ \ge \ \frac{1-\sigma^6}{1-\sigma^2} \ = \ 1+\sigma^2+\sigma^4 \ .$$

Therefore the claim follows once we show

$$X \ \le \ 4+\frac{31}{5}\,\sigma^2\left(1+\sigma^2\right) \ . \tag{67}$$

As in the proof of Lemma A.1, $a_0 = \frac{1+u^2}{2}$ for $\sigma \le \sigma^*$ and $a_0 = u^{2/3} = \sigma^2$ for $\sigma \ge \sigma^*$, where $\sigma^* := \sqrt{\frac{\sqrt{5}-1}{2}}$. We will use the bounds $(\sigma^*)^3 < 1/2$ and $\sigma^* > 3/4$ established in that proof.

Case $\sigma \le \sigma^*$. Here $a_0+u-2u^2 = \frac{(1+3u)(1-u)}{2}$ and $a_0^2-u^2 = \frac{(1-u)^2(1+u)^2}{4}$, so

$$X \ \le \ \frac{4(1+3u)}{(1-u)(1+u)^2} = \frac{4(1+3u)}{1+u-u^2-u^3} \ ,$$

$$X - 4 \;\le\; \frac{4u\left(2+u+u^2\right)}{1+u-u^2-u^3} \;\; .$$

At $\sigma = 0$ the right-hand side vanishes and (67) holds with equality, so assume $\sigma > 0$. Substituting $u = \sigma^3$ and dividing (67) by $\frac{\sigma^2}{5} > 0$, it suffices to prove

$$\begin{aligned} &20\,\sigma\left(2+\sigma^3+\sigma^6\right) \;\le\; 31\left(1+\sigma^2\right)\left(1+\sigma^3-\sigma^6-\sigma^9\right) \\ &\text{for } 0 < \sigma \le \sigma^* \;\; . \end{aligned}$$

Since $\sigma^3 < 1/2$ and $\sigma^6 < 1/4$, the left-hand side is at most

$$20\sigma\left(2+\frac{1}{2}+\frac{1}{4}\right) = 55\sigma \;\; .$$

For the right-hand side, put $w := \sigma^3 \in [0, 1/2]$. Then $1 - w - w^2 \ge 1/4$, so

$$1 + w - w^2 - w^3 = 1 + w(1 - w - w^2) \ge 1 \;\; .$$

The right-hand side is therefore at least $31(1+\sigma^2)$. This exceeds 55σ, since

$$31(1+\sigma^2) - 55\sigma = 31(1-\sigma)^2 + 7\sigma > 0 \;\; .$$

Case $\sigma \ge \sigma^*$. Here $a_0 = \sigma^2$, so $a_0^2 - u^2 = \sigma^4(1-\sigma^2)$ and $a_0 + u - 2u^2 = \sigma^2(1+\sigma-2\sigma^4)$, whence

$$X \;\le\; \frac{2\left(1+\sigma-2\sigma^4\right)}{\sigma^2\left(1-\sigma^2\right)} \;\; .$$

Multiplying (67) by $5\sigma^2(1-\sigma^2) > 0$, it suffices to prove

$$10\left(1+\sigma-2\sigma^4\right) \;\le\; \sigma^2\left(1-\sigma^2\right)\left(20+31\sigma^2+31\sigma^4\right) \;\; .$$

Using $1+\sigma-2\sigma^4 = (1-\sigma)\left(1+2\sigma+2\sigma^2+2\sigma^3\right)$ and $1-\sigma^2 = (1-\sigma)(1+\sigma)$ and dividing by $1-\sigma > 0$, this is

$$10\left(1+2\sigma+2\sigma^2+2\sigma^3\right) \;\le\; \sigma^2(1+\sigma)\left(20+31\sigma^2+31\sigma^4\right) \;\; ,$$

that is, $P(\sigma) \ge 0$ for

$$P(\sigma) \;:=\; 31\left(\sigma^4+\sigma^5+\sigma^6+\sigma^7\right) - 20\sigma - 10 \;\; .$$

The derivative

$$P'(\sigma) = 31\left(4\sigma^3+5\sigma^4+6\sigma^5+7\sigma^6\right) - 20$$

is increasing for $\sigma \ge 0$, and $P'(1/2) = 921/64 > 0$. Thus P is increasing on $[1/2, 1]$. Since $\sigma^* > 3/4$, we obtain

$$P(\sigma) \ge P(3/4) = \frac{29825}{16384} > 0 \qquad (\sigma^* \le \sigma < 1) \;\; .$$

The case $\sigma = 1$, i.e. $\delta_{\mathcal{C}} = 0$, is excluded since $\delta_{\mathcal{C}} > 0$. □

C A simpler strong converse counterexample for binary fields

We give an alternative proof of the characteristic-two case of Theorem 10.1. The construction starts from binary BCH codes with large decoding lists and an explicit minimum-weight word. Random parity checks conditioned on preserving that word raise a fixed generalized weight while retaining a large list. A random linear amplification then adjusts the distance and rate while preserving both properties. We use the definitions and the scalar-extension and MCA lemmas from Section 10.1 throughout.

Theorem C.1. *For parameters as in* (42)*, the rate, distance, list-size and MCA conclusions* (44) *and* (45) *of Theorem 10.1 with $\ell = 2$ can also be obtained from binary BCH codes by expurgation and linear amplification.*

C.1 Generalized Hamming weight amplification

We now show how to remove k-dimensional subspaces with small support from a binary linear code by imposing additional linear constraints that preserve a chosen minimum-weight word. The resulting subcode has the same minimum distance, loses at most the number of imposed constraints in dimension, and has a large k-th generalized Hamming weight. Related probabilistic existence bounds based on counting subspaces with small support appear in [ZLD13]; the constraints here are conditioned to preserve the chosen word.

Lemma C.2 (Generalized Hamming weight amplification)**.** *Let $\mathcal{C}_0 \subseteq \mathbb{F}_2^N$ be a linear code of minimum distance D, let $\mathbf{b}_* \in \mathcal{C}_0$ have weight D, and suppose some ball of absolute radius P contains at least 2^ℓ words of $\mathcal{C}_0$. Let $\sigma \geq 0$, $k \geq 2$ and $0 \leq W \leq N$ be integers with $\dim \mathcal{C}_0 - \sigma \geq k$. If*

$$2^{-\sigma(k-1)} \sum_{j=0}^{W} \binom{N}{j} 2^{kj} \;<\; 1 \;\;, \tag{68}$$

then there is a subcode $\mathcal{C} \leq \mathcal{C}_0$ with

$$\begin{aligned} &\mathbf{b}_* \in \mathcal{C} \;\;, \\ &\Delta(\mathcal{C}) = D \;\;, \\ &\dim \mathcal{C} \geq \dim \mathcal{C}_0 - \sigma \;\;, \\ &d_k(\mathcal{C}) > W \;\;, \end{aligned} \tag{69}$$

and some ball of radius P contains at least $2^{\ell-\sigma}$ words of $\mathcal{C}$.

Proof. Let $\phi_1, \ldots, \phi_\sigma$ be independent and uniform among the linear functionals on $\mathbb{F}_2^N$ vanishing on $\mathbf{b}_*$, and set $\mathcal{C} := \mathcal{C}_0 \cap \bigcap_i \ker \phi_i$.

Fix a k-dimensional $V \leq \mathcal{C}_0$. Each ϕ_i is uniform on the dual of $\mathbb{F}_2^N / \langle \mathbf{b}_* \rangle$, and it vanishes on V precisely when it vanishes on $V + \langle \mathbf{b}_* \rangle$, so

$$\Pr\left[\phi_i|_V = 0\right] = 2^{-\dim((V+\langle \mathbf{b}_* \rangle)/\langle \mathbf{b}_* \rangle)} \;\leq\; 2^{-(k-1)} \;\;,$$

since $\dim\left((V + \langle \mathbf{b}_* \rangle)/\langle \mathbf{b}_* \rangle\right) = \dim V - \dim\left(V \cap \langle \mathbf{b}_* \rangle\right) \geq k - 1$. By independence, $\Pr\left[V \leq \mathcal{C}\right] \leq 2^{-\sigma(k-1)}$.

The k-dimensional subspaces with $|\mathsf{supp}(V)| \leq W$ number at most $\sum_{j=0}^{W} \binom{N}{j} 2^{kj}$: such a V lies in $\mathbb{F}_2^U$ for some U with $|U| = j \leq W$, and is spanned by an ordered k-tuple of vectors of $\mathbb{F}_2^U$. A union bound and (68) therefore give a choice of the ϕ_i for which no such subspace survives, i.e. $d_k(\mathcal{C}) > W$.

The remaining assertions in (69) are immediate: $\mathbf{b}_* \in \mathcal{C}$ by construction, so $\Delta(\mathcal{C}) = D$ because $\mathcal{C} \le \mathcal{C}_0$; and $\dim \mathcal{C} \ge \dim \mathcal{C}_0 - \sigma$ because σ checks were imposed. Finally, the words of $\mathcal{C}_0$ in the given ball meet at most 2^σ cosets of $\mathcal{C}$ in $\mathcal{C}_0$, so some coset contains at least $2^{\ell-\sigma}$ of them; translating that coset into $\mathcal{C}$, and the center of the ball by the same vector, preserves all distances. □

C.2 Binary codes with large lists and large generalized weights

Fix $0 < \alpha < 1$ and an integer

$$b > \max\left\{2,\ \frac{1+\alpha}{\alpha}\right\}\ . \tag{70}$$

For integers $s \to \infty$ put

$$\begin{aligned} m &:= bs\ , \\ N &:= 2^m - 1\ , \\ r &:= 2^{s-1} - 1\ , \\ D &:= 2r + 1 = 2^s - 1\ , \\ P &:= \lfloor (1+\alpha) r \rfloor\ , \end{aligned} \tag{71}$$

index the coordinates by $\mathbb{F}_{2^m}^\times$, and let

$$\mathcal{C}_0 := \left\{ \mathbf{c} \in \mathbb{F}_2^N\ :\ \sum_{\xi \in \mathbb{F}_{2^m}^\times} \mathbf{c}_\xi\, \xi^j = 0 \ \text{ for odd } j \in \{1, 3, \dots, 2r-1\} \right\}\ . \tag{72}$$

The starting family $\mathcal{C}_0$ is the classical primitive narrow-sense binary BCH family, with length $N = 2^m - 1$ and designed distance $D = 2r + 1$ [Hoc59; BRC60].

Lemma C.3. *There is a fixed $\eta > 0$ such that, for all sufficiently large s, the code $\mathcal{C}_0$ has dimension at least $N - mr$, minimum distance exactly D attained by an explicit codeword $\mathbf{b}_*$, and a ball of radius P containing at least $2^{\eta r m}$ codewords.*

Proof. Each of the r checks in (72) is an $\mathbb{F}_2$-linear map from $\mathbb{F}_2^N$ to $\mathbb{F}_{2^m}$, hence imposes at most m binary constraints. Thus $\dim \mathcal{C}_0 \ge N - mr$.

To obtain the checks at even exponents, fix $\mathbf{c} \in \mathcal{C}_0$ and define

$$S_j(\mathbf{c}) := \sum_{\xi \in \mathbb{F}_{2^m}^\times} \mathbf{c}_\xi\, \xi^j$$

for $j \ge 1$. The Frobenius map $z \mapsto z^2$ is additive in characteristic two, and $\mathbf{c}_\xi^2 = \mathbf{c}_\xi$ because $\mathbf{c}_\xi \in \mathbb{F}_2$. Consequently,

$$S_j(\mathbf{c})^2 = \sum_{\xi \in \mathbb{F}_{2^m}^\times} \mathbf{c}_\xi^2\, \xi^{2j} = S_{2j}(\mathbf{c})$$

and, by iteration, for every $a \ge 0$,

$$S_j(\mathbf{c})^{2^a} = S_{2^a j}(\mathbf{c})\ .$$

Every integer $t \in \{1, \dots, 2r\}$ can be written uniquely as $t = 2^a j$ with $a \geq 0$ and j odd. Since $j \leq t \leq 2r$ and j is odd, $j \leq 2r - 1$, so $S_j(\mathbf{c}) = 0$ is one of the defining checks. Repeated squaring therefore gives

$$S_t(\mathbf{c}) = S_j(\mathbf{c})^{2^a} = 0$$

for every $1 \leq t \leq 2r$.

If a nonzero codeword had support $\{\xi_1, \dots, \xi_w\}$ with $w \leq 2r$, the first w of these checks would give

$$\left(\xi_i^j\right)_{1 \leq j, i \leq w} \begin{pmatrix} \mathbf{c}_{\xi_1} \\ \vdots \\ \mathbf{c}_{\xi_w} \end{pmatrix} = 0 \ .$$

This matrix is invertible: its determinant is $(\prod_{i=1}^{w} \xi_i) \prod_{1 \leq i < l \leq w} (\xi_l - \xi_i) \neq 0$, since the ξ_i are distinct and nonzero. This contradicts the nonzero coefficient vector, so $\Delta(\mathcal{C}_0) \geq 2r + 1 = D$.

Since $s \mid m$, the group $\mathbb{F}_{2^s}^{\times}$ sits inside $\mathbb{F}_{2^m}^{\times}$ and has order D. Let $\mathbf{b}_*$ be its indicator, of weight D. For $1 \leq j < D$ the character sum $\sum_{\xi \in \mathbb{F}_{2^s}^{\times}} \xi^j$ vanishes, so $\mathbf{b}_* \in \mathcal{C}_0$ and $\Delta(\mathcal{C}_0) = D$.

Averaging over binary centers, some $w \in \mathbb{F}_2^N$ has at least $2^{-mr}\binom{N}{P}$ codewords of $\mathcal{C}_0$ at distance exactly P. Using $\binom{N}{P} \geq (N/P)^P$ together with $\log_2 N = m + o(1)$ and $\log_2 P = m/b + O(1)$,

$$\log_2 \left(2^{-mr}\binom{N}{P}\right) \ \geq \ -mr + P \log_2(N/P) = \Big((1+\alpha)\left(1 - \tfrac{1}{b}\right) - 1 + o(1)\Big) rm \ ,$$

and the coefficient $(1 + \alpha)(1 - 1/b) - 1$ is positive precisely by the choice (70) of b. Any η below it works for large s. □

Lemma C.4. *For every fixed $T > 1$, there exist a fixed integer $k \geq 2$ and subcodes $\mathcal{C} \leq \mathcal{C}_0$ with*

$$\begin{aligned} &\dim \mathcal{C} = N - O(rm) \ , \\ &\Delta(\mathcal{C}) = D \ , \\ &d_k(\mathcal{C}) > TD \ , \end{aligned} \tag{73}$$

still having a list of at least $2^{\eta rm/2}$ codewords in a ball of radius P.

Proof. Apply Lemma C.2 to $\mathcal{C}_0$ and $\mathbf{b}_*$ with

$$\begin{aligned} &\sigma := \lfloor \eta rm/2 \rfloor \ , \\ &W := \lfloor TD \rfloor \ , \\ &k := \text{any fixed integer with } k - 1 > 4T/\eta \ . \end{aligned}$$

For fixed T and k, and using $W = \Theta(r)$ and $\log_2 N = m + o(1)$,

$$\log_2 \left(\sum_{j=0}^{W} \binom{N}{j} 2^{kj}\right) \ \leq \ \log_2(W+1) + W \log_2 N + kW = (2T + o(1))\, rm \ ,$$

whereas $\sigma(k-1) = (\eta(k-1)/2 + o(1))\, rm$ and $\eta(k-1)/2 > 2T$. Hence (68) holds for all sufficiently large s. Moreover $rm = o(N)$, so $\dim \mathcal{C}_0 - \sigma \geq N - mr - \eta rm/2 \geq k$ for large s, and the dimension hypothesis of Lemma C.2 holds as well.

Lemma C.2 now supplies $\mathcal{C} \leq \mathcal{C}_0$ with $\mathbf{b}_* \in \mathcal{C}$, $\Delta(\mathcal{C}) = D$, $\dim \mathcal{C} \geq N - mr - \sigma = N - O(rm)$, and $d_k(\mathcal{C}) > W = \lfloor TD \rfloor$; since generalized weights are integers, $d_k(\mathcal{C}) > TD$. The surviving list has size at least $2^{\eta rm - \sigma} \geq 2^{\eta rm/2}$. □

C.3 Distance and rate amplification

Fix a power of two q, write $a := 1 - 1/q$, and assume $0 < d < a$. Let $\mathcal{C}$ be as in Lemma C.4 and set

$$\begin{aligned} z &:= 1 - \frac{d}{a} \in (0,1) \ , \\ \theta &:= 1 - z^{1/D} \ , \\ \mathcal{C}^{(q)} &:= \mathcal{C} \otimes_{\mathbb{F}_2} \mathbb{F}_q \ , \\ n &:= \lceil cN \rceil \ . \end{aligned} \tag{74}$$

Let $A \in \mathbb{F}_q^{n \times N}$ be random: each entry is independently *activated* with probability θ, an activated entry is assigned a uniform element of $\mathbb{F}_q$, and an inactive entry is set to zero. Put

$$\begin{aligned} &F(w) := a\,(1 - (1-\theta)^w) = a\left(1 - z^{w/D}\right) \\ &(w \geq 0) \ , \end{aligned} \tag{75}$$

an increasing function of w with $F(D) = d$.

Related distance and rate amplification by random XOR and linear row operations appears in [RW15]. Here we use the activated-entry distribution above and also control the support of subspaces.

Lemma C.5 (One row)**.** *Let ϕ be a row of A. For a fixed $\mathbf{v} \in \mathbb{F}_q^N$ of weight w, $\Pr[\phi \cdot \mathbf{v} \neq 0] = F(w)$. For a fixed subspace $V \leq \mathbb{F}_q^N$ with $|\mathsf{supp}(V)| = w$, $\Pr[\phi|_V \neq 0] \geq F(w)$.*

Proof. Condition on the set $\mathcal{A}$ of activated entries of ϕ. If $\mathcal{A}$ misses $\mathsf{supp}(\mathbf{v})$ then $\phi \cdot \mathbf{v} = 0$; otherwise $\phi \cdot \mathbf{v}$ is a sum containing at least one independent uniform element of $\mathbb{F}_q$, hence is itself uniform on $\mathbb{F}_q$ and nonzero with probability a. Since $\Pr[\mathcal{A} \cap \mathsf{supp}(\mathbf{v}) \neq \varnothing] = 1 - (1-\theta)^w$, the first claim follows.

For the second, condition on $\mathcal{A}$ with $\mathcal{A} \cap \mathsf{supp}(V) \neq \varnothing$. The functional $\phi|_V$ is then the image of a uniform vector of $\mathbb{F}_q^{\mathcal{A} \cap \mathsf{supp}(V)}$ under the linear map sending the i-th basis vector to $e_i^\vee|_V$, so it is uniform on a subspace $\mathcal{S} \leq V^\vee$; and $\mathcal{S} \neq 0$ because $e_i^\vee|_V \neq 0$ for every $i \in \mathsf{supp}(V)$. Hence $\Pr[\phi|_V \neq 0 \mid \mathcal{A}] = 1 - 1/|\mathcal{S}| \geq a$, and we conclude as before. □

Lemma C.6 (Amplification)**.** *Fix $T > 1$, and let $(\mathcal{C}_s)_{s \geq s_0}$ be a family of binary codes supplied by Lemma C.4, with α, b, η and k fixed, and with*

$$\begin{aligned} m_s &:= bs \ , \\ N_s &:= 2^{bs} - 1 \ , \\ r_s &:= 2^{s-1} - 1 \ , \\ \text{and } D_s &:= 2^s - 1 \ . \end{aligned}$$

Fix q and d as in (74)*, and put $H := a(1 - z^T)$. There exists a constant $c_0 > 0$ such that, for every fixed $c \geq c_0$, setting $n_s := \lceil cN_s \rceil$, one can choose matrices $A_s \in \mathbb{F}_q^{n_s \times N_s}$ for all sufficiently large integers s so that the codes*

$$\widehat{\mathcal{C}}_s := A_s\,(\mathcal{C}_s \otimes_{\mathbb{F}_2} \mathbb{F}_q) \subseteq \mathbb{F}_q^{n_s}$$

satisfy

$$\dim_{\mathbb{F}_q} \widehat{\mathcal{C}}_s = \dim_{\mathbb{F}_2} \mathcal{C}_s \text{ for all sufficiently large } s \ ,$$

$$\lim_{s\to\infty} \delta(\widehat{\mathcal{C}}_s) = d \ ,$$
$$\text{and } \liminf_{s\to\infty} \frac{d_k\left(\widehat{\mathcal{C}}_s\right)}{n_s} \geq H \ . \tag{76}$$

Moreover, there is a sequence $\varepsilon_s \geq 0$ with $\varepsilon_s \to 0$ as $s \to \infty$ such that, for every sufficiently large s, some ball in $\mathbb{F}_q^{n_s}$ of relative radius

$$p_* + \varepsilon_s \tag{77}$$

contains at least $2^{\eta r_s m_s/2}$ distinct codewords of $\widehat{\mathcal{C}}_s$. Here $p_* := a\left(1 - z^{(1+\alpha)/2}\right)$.

Proof. We suppress the subscript s on the codes, matrices and parameters throughout the proof; all limits are taken as $s \to \infty$ with α, b, η, k, T, q, d and c fixed. Fix $T' > T$, and put $L_N := \lfloor T'D \rfloor$. Since $D = 2^s - 1$ and $N = 2^{bs} - 1$ with fixed $b > 2$, we have $D \log N = o(N)$. Define

$$\tau_N := \left(\frac{D \log N}{N}\right)^{1/4} \ .$$

Then $\tau_N \to 0$ and

$$\frac{N\tau_N^2}{D \log N} = \sqrt{\frac{N}{D \log N}} \longrightarrow \infty \ .$$

Concentration for a fixed vector or subspace. Write $\phi_1, \ldots, \phi_n$ for the independent rows of A. For a fixed vector $\mathbf{v}$ of weight w, the indicators $\mathbb{1}_{\{\phi_i \cdot \mathbf{v} \neq 0\}}$ are independent Bernoulli variables of mean $F(w)$ by Lemma C.5, and their sum is $\mathsf{wt}(A\mathbf{v})$. For a fixed subspace V with $|\mathsf{supp}(V)| = w$, the indicators $\mathbb{1}_{\{\phi_i|_V \neq 0\}}$ are independent Bernoulli variables of a common mean $p_V \geq F(w)$, and their sum is $|\mathsf{supp}(AV)|$. Thus, for every $t > 0$, Hoeffding's inequality gives

$$\Pr\left[\left|\frac{\mathsf{wt}(A\mathbf{v})}{n} - F(w)\right| > t\right] \leq 2\exp(-2nt^2) \ ,$$
$$\text{and } \Pr\left[\frac{|\mathsf{supp}(AV)|}{n} < F(w) - t\right] \leq \Pr\left[\frac{|\mathsf{supp}(AV)|}{n} < p_V - t\right] \leq \exp(-2nt^2) \ .$$

Small support. The number of vectors of weight at most L_N is at most

$$\sum_{j=0}^{L_N} \binom{N}{j} (q-1)^j \leq (L_N + 1) N^{L_N} q^{L_N} \ .$$

For a fixed coordinate set U of size j, every k-dimensional subspace supported on U has an ordered basis in $(\mathbb{F}_q^U)^k$, so there are at most q^{kj} such subspaces. Consequently, the number of k-dimensional subspaces with support size at most L_N is at most

$$\sum_{j=0}^{L_N} \binom{N}{j} q^{kj} \leq (L_N + 1) N^{L_N} q^{kL_N} \ .$$

Applying the preceding concentration bounds with $t = \tau_N$ and taking a union bound, the probability that either a vector of weight at most L_N or a k-dimensional subspace of support size at most L_N violates its corresponding estimate is at most

$$\begin{aligned} 3(L_N+1)N^{L_N}q^{kL_N}\exp(-2n\tau_N^2) &= \exp\left(\log 3 + \log(L_N+1) + L_N\log N + kL_N\log q - 2n\tau_N^2\right) \\ &\le \exp\left(O(D\log N) - 2c\sqrt{ND\log N}\right) \longrightarrow 0 \ . \end{aligned}$$

Here q, k and T' are fixed, $L_N \le T'D$, and $n \ge cN$. Hence, with probability tending to one, simultaneously

$$\left|\frac{\mathsf{wt}(A\mathbf{v})}{n} - F(\mathsf{wt}(\mathbf{v}))\right| \le \tau_N$$

for every $\mathbf{v}$ of weight at most L_N, and

$$\frac{|\mathsf{supp}(AV)|}{n} \ge F(|\mathsf{supp}(V)|) - \tau_N$$

for every k-dimensional V with support size at most L_N.

Large support. Put

$$\begin{aligned} \mu_* &:= F(T'D) = a(1-z^{T'}) \ , \\ \varepsilon_d &:= \mu_* - d > 0 \ , \\ \text{and } \varepsilon_H &:= \mu_* - H > 0 \ . \end{aligned}$$

The strict inequalities follow from $T' > T > 1$ and $0 < z < 1$. If $\mathsf{wt}(\mathbf{v}) > L_N$, then $\mathsf{wt}(\mathbf{v}) > T'D$ because the weight is integral, and therefore $F(\mathsf{wt}(\mathbf{v})) \ge \mu_*$. The one-sided concentration bound gives

$$\Pr\left[\frac{\mathsf{wt}(A\mathbf{v})}{n} \le d\right] \le \exp(-2n\varepsilon_d^2) \ .$$

Likewise, if $|\mathsf{supp}(V)| > L_N$, then $p_V \ge \mu_*$, and

$$\Pr\left[\frac{|\mathsf{supp}(AV)|}{n} \le H\right] \le \exp(-2n\varepsilon_H^2) \ .$$

There are at most q^N vectors and at most q^{kN} k-dimensional subspaces in $\mathbb{F}_q^N$, the latter again by counting ordered bases. Set $\varepsilon := \min\{\varepsilon_d, \varepsilon_H\}$ and choose any fixed

$$c > \frac{k\log q + 1}{2\varepsilon^2} \ .$$

The probability that any of these large-support vectors has image relative weight at most d, or any of these large-support subspaces has image support fraction at most H, is bounded by

$$\begin{aligned} q^N\exp(-2n\varepsilon_d^2) + q^{kN}\exp(-2n\varepsilon_H^2) &\le 2\exp\left((k\log q - 2c\varepsilon^2)N\right) \\ &\le 2e^{-N} \longrightarrow 0 \ . \end{aligned}$$

Together with the small-support estimate, this proves that all four conclusions hold simultaneously with probability tending to one. For each sufficiently large s, fix a matrix A satisfying them.

Injectivity, distance, and generalized weight. By Lemmas C.4 and 10.4, every nonzero $\mathbf{v} \in \mathcal{C}^{(q)}$ has weight at least D. If its weight is at most L_N, the small-support estimate and monotonicity of F imply

$$\frac{\mathsf{wt}(A\mathbf{v})}{n} \geq F(\mathsf{wt}(\mathbf{v})) - \tau_N \geq F(D) - \tau_N = d - \tau_N \ .$$

If its weight is larger than L_N, the large-support estimate gives $\mathsf{wt}(A\mathbf{v})/n > d$. Since $\tau_N < d$ for all sufficiently large s, every nonzero $\mathbf{v} \in \mathcal{C}^{(q)}$ has nonzero image. Thus A is injective on $\mathcal{C}^{(q)}$, and

$$\begin{aligned} \dim \widehat{\mathcal{C}} &= \dim \mathcal{C}^{(q)} = \dim \mathcal{C} \ , \\ \text{and } \delta(\widehat{\mathcal{C}}) &\geq d - \tau_N \ . \end{aligned}$$

The codeword $\mathbf{b}_*$ has weight exactly $D \leq L_N$, so the small-support estimate also gives

$$\delta(\widehat{\mathcal{C}}) \leq \frac{\mathsf{wt}(A\mathbf{b}_*)}{n} \leq F(D) + \tau_N = d + \tau_N \ .$$

It follows that $\delta(\widehat{\mathcal{C}}) \to d$.

Every k-dimensional subspace $\widehat{V} \leq \widehat{\mathcal{C}}$ is the image AV of a k-dimensional $V \leq \mathcal{C}^{(q)}$, by injectivity. Scalar extension preserves generalized weights, so $w := |\mathsf{supp}(V)| > TD$. If $w \leq L_N$, then

$$\frac{|\mathsf{supp}(\widehat{V})|}{n} \geq F(w) - \tau_N \geq F(TD) - \tau_N = H - \tau_N \ .$$

If $w > L_N$, the large-support estimate gives $|\mathsf{supp}(\widehat{V})|/n > H$. Taking the minimum over all such $\widehat{V}$ yields

$$\frac{d_k\left(\widehat{\mathcal{C}}\right)}{n} \geq H - \tau_N \ ,$$

and hence the asserted lower bound on the limit inferior.

Preserving the decoding list. Let w be the binary center and $\mathcal{L} \subseteq \mathcal{C}$ the list from Lemma C.4. For each $\mathbf{c} \in \mathcal{L}$, we have $\mathsf{wt}(w - \mathbf{c}) \leq P < D \leq L_N$. The small-support estimate applies to this error vector, even though it need not lie in $\mathcal{C}^{(q)}$, and gives

$$\begin{aligned} \frac{\mathsf{wt}(Aw - A\mathbf{c})}{n} &\leq F(\mathsf{wt}(w - \mathbf{c})) + \tau_N \\ &\leq F(P) + \tau_N \ . \end{aligned}$$

Since $P/D \to (1+\alpha)/2$, the radius on the right satisfies

$$F(P) + \tau_N = a(1 - z^{P/D}) + \tau_N \longrightarrow a(1 - z^{(1+\alpha)/2}) = p_* \ .$$

Injectivity of A on $\mathcal{C}^{(q)}$ ensures that the images of the words in $\mathcal{L}$ are distinct. Taking $\varepsilon_s := |F(P) - p_*| + \tau_N$, we have $\varepsilon_s \geq 0$, $\varepsilon_s \to 0$, and $F(P) + \tau_N \leq p_* + \varepsilon_s$. The ball centered at Aw with relative radius $p_* + \varepsilon_s$ therefore contains at least $|\mathcal{L}| \geq 2^{\eta r m/2}$ words of $\widehat{\mathcal{C}}$. □

C.4 Proof of Theorem C.1

Proof. Fix d, p, γ as in (42). The constants are chosen in the following order.

The field. Choose a power of two q large enough that, with $a = 1 - 1/q$ and $z = 1 - d/a$,

$$\begin{aligned} & d < a \ , \\ & \Psi(\gamma) < a \ , \\ & a\left(1 - \sqrt{z}\right) < p \ . \end{aligned} \tag{78}$$

All three hold for every large q: the first because $a \to 1 > d$; the second because $\Psi(\gamma) < 1$, as $\gamma < 1$; and the third because $z \to 1 - d$, so that $a(1 - \sqrt{z}) \to J_2(d) < p$.

The list radius. Choose $\alpha \in (0, 1)$ small enough that

$$p_* = a\left(1 - z^{(1+\alpha)/2}\right) < p \ , \tag{79}$$

which is possible since the left-hand side tends to $a(1 - \sqrt{z}) < p$ as $\alpha \to 0$.

The generalized weight. Choose $T > 1$ large enough that

$$\Psi(\gamma) \ < \ H = a\left(1 - z^T\right) \ , \tag{80}$$

which is possible since $H \to a > \Psi(\gamma)$ as $T \to \infty$, by (78).

The codes. With α fixed, choose b as in (70) and let $\eta > 0$ be as in Lemma C.3; with T fixed, let k and the subcodes $\mathcal{C}$ be as in Lemma C.4. Apply Lemma C.6 with a sufficiently large fixed c, and call the resulting codes $\mathcal{C}_n \subseteq \mathbb{F}_q^n$, $n = \lceil cN \rceil$.

Rate and distance. By (76) and Lemma C.4, $\dim \mathcal{C}_n = \dim \mathcal{C} = N - O(rm)$, and $rm = o(N)$, so the rate tends to $R = 1/c > 0$; and $\delta(\mathcal{C}_n) \to d$ by (76). This proves (44).

Preserving the decoding list. By (77) and (79), for all sufficiently large n the ball produced by Lemma C.6 has relative radius below p, so

$$L_{\mathcal{C}_n}(p) \ \geq \ 2^{\eta r m/2} \ \geq \ 2^{c_0 n^{1/b} \log_2 n} \tag{81}$$

for a constant $c_0 > 0$: indeed $r = \Theta(2^s)$, $m = bs$ and $n = \Theta(2^{bs})$, so $rm = \Theta\left(n^{1/b} \log_2 n\right)$.

The gap. By (76) and (80) there is a fixed $\zeta_0 > 0$ with

$$1 - \frac{n\,\Psi(\gamma)}{d_k(\mathcal{C}_n)} \ \geq \ \zeta_0 \tag{82}$$

for all sufficiently large n.

Scalar extension. Let $\mathbb{K}/\mathbb{F}_q$ be any finite extension and $\mathcal{C}_{n,\mathbb{K}} = \mathcal{C}_n \otimes_{\mathbb{F}_q} \mathbb{K}$. By Lemma 10.4 the dimension and the minimum distance are unchanged, so (44) persists; and the list of (81), whose center lies in $\mathbb{F}_q^n$, is still a list in the larger code, so the first bound in (45) persists as well. Finally Lemma 10.5 together with (82) gives

$$\epsilon_{\mathsf{MCA}}^{\mathcal{C}_{n,\mathbb{K}}}(\gamma) \ \leq \ \frac{\mathsf{T}_k(n,q) + 3}{\zeta_0\,|\mathbb{K}|} \ \leq \ \frac{C_0 n^k}{|\mathbb{K}|}$$

for a fixed C_0, since $\mathsf{T}_k(n,q) \leq k q^{k-1} n^k = O(n^k)$ with k and q fixed. Taking $\beta = 1/b$ completes the proof. □

Acknowledgements

The author is supported by the Ethereum Foundation and Polygon Labs. The author thanks Alessandro Chiesa for useful discussions and guidance in this project.

References

[ACFY24] Gal Arnon, Alessandro Chiesa, Giacomo Fenzi, and Eylon Yogev. "STIR: Reed–Solomon Proximity Testing with Fewer Queries". In: *Advances in Cryptology – CRYPTO 2024*. Vol. 14929. Lecture Notes in Computer Science. 2024, pp. 380–413.

[ACFY25] Gal Arnon, Alessandro Chiesa, Giacomo Fenzi, and Eylon Yogev. "WHIR: Reed–Solomon Proximity Testing with Super-Fast Verification". In: *Advances in Cryptology – EUROCRYPT 2025*. Vol. 15604. Lecture Notes in Computer Science. 2025, pp. 214–243.

[AGG26] Omar Alrabiah, Rohan Goyal, and Venkatesan Guruswami. *Explicit Reed–Solomon Subspace Designs: Optimal List and Line Decoding*. Electronic Colloquium on Computational Complexity, Report 2026/262. 2026. URL: https://eccc.weizmann.ac.il/report/2026/262/.

[AGL24] Omar Alrabiah, Venkatesan Guruswami, and Ray Li. "Randomly Punctured Reed–Solomon Codes Achieve List-Decoding Capacity over Linear-Sized Fields". In: *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*. STOC '24. 2024, pp. 1458–1469.

[AHIV17] Scott Ames, Carmit Hazay, Yuval Ishai, and Muthuramakrishnan Venkitasubramaniam. "Ligero: Lightweight Sublinear Arguments Without a Trusted Setup". In: *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*. CCS '17. 2017, pp. 2087–2104.

[AK11] Per Austrin and Subhash Khot. "A Simple Deterministic Reduction for the Gap Minimum Distance of Code Problem". In: *Automata, Languages and Programming – ICALP 2011*. Vol. 6755. Lecture Notes in Computer Science. 2011, pp. 474–485.

[AK24] Benny Applebaum and Eliran Kachlon. "Conflict Checkable and Decodable Codes and Their Applications". In: *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms*. SODA '24. Full version available at the linked ECCC record. 2024, pp. 1379–1424. URL: https://eccc.weizmann.ac.il/report/2023/062/.

[BCGM26] Sarah Bordage, Alessandro Chiesa, Ziyi Guan, and Ignacio Manzur. "All Polynomial Generators Preserve Distance with Mutual Correlated Agreement". In: *41st Computational Complexity Conference (CCC 2026)*. Vol. 383. Leibniz International Proceedings in Informatics (LIPIcs). 2026, 24:1–24:18. URL: https://eprint.iacr.org/2025/2051.

[BCHKS26] Eli Ben-Sasson, Dan Carmon, Ulrich Haböck, Swastik Kopparty, and Shubhangi Saraf. "On Proximity Gaps of Reed–Solomon Codes". In: *Proceedings of the 58th Annual ACM Symposium on Theory of Computing*. STOC '26. 2026, pp. 1157–1167.

[BCIKS20] Eli Ben-Sasson, Dan Carmon, Yuval Ishai, Swastik Kopparty, and Shubhangi Saraf. "Proximity Gaps for Reed–Solomon Codes". In: *Proceedings of the 61st Annual IEEE Symposium on Foundations of Computer Science*. FOCS '20. 2020, pp. 900–909.

[BCPZZ26] Joshua Brakensiek, Yeyuan Chen, Aaron Putterman, Zihan Zhang, and Kai Zhe Zheng. *Algorithmic List Decoding of Reed–Solomon Codes up to Capacity*. Electronic Colloquium on Computational Complexity, Report 2026/164, Revision 1. 2026. URL: https://eccc.weizmann.ac.il/report/2026/164/.

[BGKS20] Eli Ben-Sasson, Lior Goldberg, Swastik Kopparty, and Shubhangi Saraf. “DEEP-FRI: Sampling Outside the Box Improves Soundness”. In: *Proceedings of the 11th Innovations in Theoretical Computer Science Conference*. Vol. 151. Leibniz International Proceedings in Informatics (LIPIcs). 2020, 5:1–5:32. URL: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.ITCS.2020.5.

[BGM23] Joshua Brakensiek, Sivakanth Gopi, and Visu Makam. “Generic Reed–Solomon Codes Achieve List-Decoding Capacity”. In: *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*. STOC ’23. 2023, pp. 1488–1501.

[BRC60] R. C. Bose and D. K. Ray-Chaudhuri. “On a Class of Error Correcting Binary Group Codes”. In: *Information and Control* 3.1 (1960), pp. 68–79.

[Cho26a] Przemek Chojecki. *Grande Finale v4: Exact Threshold Compilers, Primitive Q, and Order-32 Reed–Solomon MCA*. Public working draft. Version of 24 July 2026. 2026. URL: https://github.com/przchojecki/rs-mca/blob/b13de8113a03f06b6fc22bbd2f289a8abcdf7e95/experimental/grande_finale.tex.

[Cho26b] Przemek Chojecki. *Shortening Bounds for Reed–Solomon MCA*. IACR Cryptology ePrint Archive, Paper 2026/1463. 2026. URL: https://eprint.iacr.org/2026/1463.

[DKTZ26] Quang Dao, Scott Duke Kominers, Justin Thaler, and Kai Zhe Zheng. *Counterexamples to Beyond-Johnson Proximity Gaps over Binary Fields*. Electronic Colloquium on Computational Complexity, Report 2026/237. 2026. URL: https://eccc.weizmann.ac.il/report/2026/237/.

[DeV14] Matt DeVos. “A Short Proof of Kneser’s Addition Theorem for Abelian Groups”. In: *Combinatorial and Additive Number Theory*. Ed. by Melvyn B. Nathanson. Vol. 101. Springer Proceedings in Mathematics & Statistics. 2014, pp. 39–41.

[ER60] Paul Erdős and Richard Rado. “Intersection Theorems for Systems of Sets”. In: *Journal of the London Mathematical Society* s1-35.1 (1960), pp. 85–90.

[Eli57] Peter Elias. “List Decoding for Noisy Channels”. In: *1957 IRE WESCON Convention Record, Part 2*. 1957, pp. 99–104. URL: https://dspace.mit.edu/handle/1721.1/4484.

[FKS22] Asaf Ferber, Matthew Kwan, and Lisa Sauermann. “List-decodability with large radius for Reed–Solomon codes”. In: *Proceedings of the 62nd Annual IEEE Symposium on Foundations of Computer Science*. FOCS ’21. 2022, pp. 720–726.

[GCXK26] Yiwen Gao, Dongliang Cai, Yang Xu, and Haibin Kan. “From List-Decodability to Proximity Gaps”. In: *IEEE Transactions on Information Theory* 72.6 (2026), pp. 3809–3819.

[GG26] Rohan Goyal and Venkatesan Guruswami. “Optimal Proximity Gaps for Subspace-Design Codes and (Random) Reed–Solomon Codes”. In: *Proceedings of the 58th Annual ACM Symposium on Theory of Computing*. STOC ’26. 2026, pp. 1558–1568.

[GGR09] Parikshit Gopalan, Venkatesan Guruswami, and Prasad Raghavendra. “List Decoding Tensor Products and Interleaved Codes”. In: *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*. STOC ’09. Full version available at the linked arXiv record. 2009, pp. 13–22. URL: https://arxiv.org/abs/0811.4395.

[GGSW26] Rohan Goyal, Venkatesan Guruswami, Yihang Sun, and Mary Wootters. *Locality of Curve-Decoding and Improved Proximity Gaps*. 2026. arXiv: 2607.08516. URL: https://arxiv.org/abs/2607.08516.

[GHK10] Venkatesan Guruswami, Johan Håstad, and Swastik Kopparty. “On the List-Decodability of Random Linear Codes”. In: *Proceedings of the 42nd ACM Symposium on Theory of Computing*. STOC ’10. 2010, pp. 409–416.

[GHSZ00] Venkatesan Guruswami, Johan Håstad, Madhu Sudan, and David Zuckerman. “Combinatorial Bounds for List Decoding”. In: *Proceedings of the 38th Annual Allerton Conference on Communication, Control, and Computing*. 2000, pp. 603–612.

[GKL24] Yiwen Gao, Haibin Kan, and Yuan Li. *Linear Proximity Gap for Linear Codes within the 1.5 Johnson Bound*. IACR Cryptology ePrint Archive, Report 2024/1810. 2024. URL: https://eprint.iacr.org/2024/1810.

[GL24] Sudhir R. Ghorpade and Rati Ludhani. "On the Minimum Distance, Minimum Weight Codewords, and the Dimension of Projective Reed–Muller Codes". In: *Advances in Mathematics of Communications* 18.2 (2024), pp. 360–382. URL: https://arxiv.org/abs/2309.10196.

[GLSTW22] Zeyu Guo, Ray Li, Chong Shangguan, Itzhak Tamo, and Mary Wootters. "Improved List-Decodability and List-Recoverability of Reed–Solomon Codes via Tree Packings: [Extended Abstract]". In: *Proceedings of the 62nd Annual IEEE Symposium on Foundations of Computer Science*. FOCS '21. 2022, pp. 708–719.

[GR06] Venkatesan Guruswami and Atri Rudra. "Explicit Capacity-Achieving List-Decodable Codes". In: *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*. STOC '06. 2006, pp. 1–10.

[GS98] Venkatesan Guruswami and Madhu Sudan. "Improved Decoding of Reed–Solomon and Algebraic-Geometric Codes". In: *Proceedings of the 39th Annual Symposium on Foundations of Computer Science*. FOCS '98. 1998, pp. 28–37.

[GYXK26] Yiwen Gao, Hong Yang, Yang Xu, and Haibin Kan. *List-Decoding Counterexamples Yield Lower Bounds on Mutual Correlated Agreement Error*. 2026. arXiv: 2607.10572. URL: https://arxiv.org/abs/2607.10572.

[GZ23] Zeyu Guo and Zihan Zhang. "Randomly Punctured Reed–Solomon Codes Achieve the List Decoding Capacity over Polynomial-Size Alphabets". In: *Proceedings of the 64th IEEE Symposium on Foundations of Computer Science*. FOCS '23. 2023, pp. 164–176.

[Gab85] E. M. Gabidulin. "Theory of Codes with Maximum Rank Distance". In: *Problems of Information Transmission* 21.1 (1985), pp. 1–12. URL: https://www.mathnet.ru/eng/ppi967.

[HKS26] Prahladh Harsha, Mrinal Kumar, and Ramprasad Saptharishi. *Reed–Solomon Codes at Capacity: Algorithmic List-Decoding and Proximity Gaps*. 2026. arXiv: 2610.08610. URL: https://arxiv.org/abs/2610.08610.

[Hab24] Ulrich Haböck. *Basefold in the List Decoding Regime*. Cryptology ePrint Archive, Paper 2024/1571. 2024. URL: https://eprint.iacr.org/2024/1571.

[Hab25] Ulrich Haböck. *A Note on Mutual Correlated Agreement for Reed–Solomon Codes*. Cryptology ePrint Archive, Paper 2025/2110. 2025. URL: https://eprint.iacr.org/2025/2110.

[Hoc59] Alexis Hocquenghem. "Codes correcteurs d'erreurs". In: *Chiffres* 2 (1959), pp. 147–156. URL: https://db.aconit.org/dbmedia_0/pdf_8/8514.pdf.

[Jer26] Fernando Granha Jeronimo. *Algorithmic List Decoding at Capacity and Optimal Proximity Gaps for Reed–Solomon Codes*. 2026. arXiv: 2609.05870. URL: https://arxiv.org/abs/2609.05870.

[Juk11] Stasys Jukna. *Extremal Combinatorics: With Applications in Computer Science*. 2nd ed. 2011.

[KTZ26] Scott Duke Kominers, Justin Thaler, and Kai Zhe Zheng. *The One-and-a-Half Johnson Bound Is Tight for Proximity Gaps of General Linear Codes*. Electronic Colloquium on Computational Complexity, Report 2026/235. 2026. URL: https://eccc.weizmann.ac.il/report/2026/235/.

[LXYZ26] Songsong Li, Chaoping Xing, Chen Yuan, and Ruiqi Zhu. *Proximity Gaps for Gabidulin Codes and Applications*. 2026. arXiv: 2609.09838. URL: https://arxiv.org/abs/2609.09838.

[RVW13] Guy N. Rothblum, Salil Vadhan, and Avi Wigderson. "Interactive Proofs of Proximity: Delegating Computation in Sublinear Time". In: *Proceedings of the 45th Annual ACM Symposium on Theory of Computing*. STOC '13. 2013, pp. 793–802.

[RW14] Atri Rudra and Mary Wootters. "Every list-decodable code for high noise has abundant near-optimal rate puncturings". In: *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*. STOC '14. 2014, pp. 764–773.

[RW15] Atri Rudra and Mary Wootters. "It'll Probably Work Out: Improved List-Decoding through Random Operations". In: *Proceedings of the 6th Innovations in Theoretical Computer Science Conference*. ITCS '15. Full version available at the linked arXiv record. 2015, pp. 287–296. URL: https://arxiv.org/abs/1408.2237.

[ST20] Chong Shangguan and Itzhak Tamo. "Combinatorial list-decoding of Reed–Solomon codes beyond the Johnson radius". In: *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*. STOC '20. 2020, pp. 538–551.

[Sud96] Madhu Sudan. "Maximum Likelihood Decoding of Reed Solomon Codes". In: *Proceedings of the 37th Annual Symposium on Foundations of Computer Science*. FOCS '96. 1996, pp. 164–172.

[Wei91] Victor K. Wei. "Generalized Hamming Weights for Linear Codes". In: *IEEE Transactions on Information Theory* 37.5 (1991), pp. 1412–1418.

[Woz58] John M. Wozencraft. *List Decoding*. Quarterly Progress Report 48. Research Laboratory of Electronics, Massachusetts Institute of Technology, 1958, pp. 90–95.

[ZLD13] Zhuojun Zhuang, Yuan Luo, and Bin Dai. "Code Constructions and Existence Bounds for Relative Generalized Hamming Weight". In: *Designs, Codes and Cryptography* 69.3 (2013), pp. 275–297.

[ZP81] V. V. Zyablov and M. S. Pinsker. "List Concatenated Decoding". In: *Problems of Information Transmission* 17.4 (1981), pp. 236–240. URL: https://www.mathnet.ru/eng/ppi1415.

[Zei26] Hadas Zeilberger. "Khatam: Proximity Gaps for Multilinear Evaluation for All Linear Codes". In: *Advances in Cryptology – CRYPTO 2026*. Vol. 16808. Lecture Notes in Computer Science. 2026, pp. 592–611. URL: https://eprint.iacr.org/2024/1843.